Almost quantum theory

Benjamin Schumacher[*] and Michael D. Westmoreland[†]

March 2, 2013

1 Introductory Remarks

1.1 Motivation

The remarkable features of quantum theory are best appreciated by comparing the theory to other possible theories—what Spekkens calls "foil" theories [1]. The most celebrated example of this approach was Bell's analysis [2], which showed that entangled quantum systems have statistical properties unlike any hypothetical local hidden variable model. More recently, there have been several efforts to give quantum theory an operational axiomatic foundation [3, 4, 5]. In these efforts, a general abstract framework is posited to describe system preparations, choices of measurement, observed results of measurement, and probabilities. Many possible theories can be expressed in the framework. The axioms embody fundamental aspects of quantum theory that uniquely identify it among them. A striking lesson of this work is that familiar quantum theory can be characterized by axioms that seem to have little to do with the traditional quantum machinery of states and observables in Hilbert space. The Hilbert space structure is "derived" from the operational axioms.

These approaches are based on two distinct concepts of generalization. First, we generalize within quantum theory to give the theory its most general form. For example, we generalize state vectors to density operators as a

[*]Department of Physics, Kenyon College. Email schumacherb@kenyon.edu
[†]Department of Mathematical Sciences, Denison University. Email westmoreland@denison.edu

description of the quantum state of a system. Second, we generalize beyond quantum theory so that we can embed it within a wider universe of possible theories. To be clear, we refer to these two processes as *development* within a theoretical framework and *extension* beyond that framework.

In this paper, we undertake these processes of development and extension, not for actual quantum theory (AQT), but for a close mathematical cousin of that theory. Modal quantum theory (MQT) [6] is a simplified toy model that reproduces many of the structural features of actual quantum theory. The underlying state space of MQT is a vector space $\mathcal{V}$ over an arbitrary field $\mathcal{F}$, which may be finite. MQT predicts, not the probabilities of the results of a measurement, but only which of those results are possible. This motivates the use of the term "modal", which in formal logic refers to operators asserting the possibility or necessity of a proposition [7]. Modal theories themselves can therefore be viewed as generalizations (extensions) of probabilistic theories.

1.2   Generalization

What is "generalization"? We begin our answer to this question with a simple example. Suppose we are devising simple substitution ciphers for English text. Each letter in the alphabet $\mathcal{A} = \{A, B, \ldots, Z\}$ is to be represented by some letter in $\mathcal{A}$. To begin with, we consider only extremely simple "transposition ciphers" in which exactly two letters are exchanged. For instance, we might exchange E and R, leaving all other letters alone.[1] An enciphered message can be decoded by applying the same transposition a second time.

To generalize this and make better ciphers, we now form compound ciphers by applying two or more transposition ciphers successively. Enciphered messages are decoded by applying the same transpositions in reverse order. Any cipher constructed out of transpositions can be described by a *permutation* of $\mathcal{A}$, an element of $S_\mathcal{A}$. Furthermore, any "permutation cipher" in $S_\mathcal{A}$ can be constructed in exactly this way, as a compound of pairwise transpositions. Thus, the concept of a permutation cipher is really a development of the original idea of a transposition.

Is further development possible? Consider the essential requirements for a "reasonable" cipher. A general substitution cipher c is a function $c : \mathcal{A} \to \mathcal{A}$. Since we need to be able to recover our plaintext correctly, it is appropriate to require as an axiom that c be a one-to-one function, so that each let-

---

[1]Such a ciphre is not vrey haed to erad.

ter in the ciphertext can be decoded in only one way. Since $\mathcal{A}$ is finite, all one-to-one functions on $\mathcal{A}$ are permutations. This means that the generalization from transpositions to permutations already encompasses all reasonable substitution ciphers (as characterized by our axiom).

To generalize further, we must extend the idea of a cipher beyond simple letter substitutions. We might apply different substitution maps to different letters, or encipher messages word-by-word. These more general ciphers will have some characteristics in common with substitution ciphers (such as the unique decipherability of an enciphered message), but will constitute a larger universe within which the substitution ciphers form a special class.

In the cipher story we can identify some general features. We begin with a basic theory based on a concept X. The process of development can involve several stages:

- *Construction.* In this stage, we devise a situation within the existing theory—that is, a situation that can be described using X—and show how this situation can be given a simpler or more natural description using X'. (*Compositions of transposition ciphers can be described as permutation ciphers.*)

- *Feasibility.* Often we are able to show that, if a situation is described by X', then it can always be given a more cumbersome description in terms of X. Informally, every instance of X' is feasible to construct from X. (*Every permutation cipher can be described as a composition of transposition ciphers.*)

- *Axiomatic characterization.* We may be led to impose one or more reasonable axioms that any situation ought to satisfy. Our development is most successful if we can establish that any "reasonable" situation (according to our axioms) can be encompassed by our generalized concept X'. (*Any uniquely decodable substitution cipher must be a permutation cipher.*)

If X' is feasible, then every instance of X' could be given a more cumbersome description in terms of X. In this case, the theory including X' is simply a development of the original one based on X. An axiomatic characterization tells us that the development is *complete*—that no further reasonable generalization of X is possible within the basic theory.

Once we have a complete development from X to X', further generalization must be an extension of the original theory.

3

- *Extension.* We can devise a broader framework Y within which the theory based on X is a special case. (*We can consider ciphers that are not based on letter-by-letter substitutions.*)

Once we have an extended framework Y, it is useful to ask what special properties the original theory may possess. Thus, we might investigate what distinguishing characteristics quantum theory has within the wider universe of probabilistic theories.

1.3 Scope of the present paper

The elementary features of modal quantum theory have been presented elsewhere [6, 8]. In the next section, we will briefly discuss the axioms for MQT, drawing the analogies between this theory and AQT. We will also discuss some of the properties of entangled states of simple systems in MQT.

Following this, we turn to a development of MQT analogous to the standard generalizations of states, measurements and dynamical evolution in AQT. Systems whose preparations are incompletely known, or which are entangled with other systems, require a more general description of their states. Measurement procedures and dynamical evolution for open systems require additional generalizations, which we will also explore. As in AQT, we can give axiomatic characterizations for these new concepts within the theory, showing that our development is, in the sense given above, complete.

To generalize further, we must embed MQT within a larger class of modal theories. We do this by analogy to the general probabilistic theories that have been used to analyze AQT. As in those theories, our modal theories are assumed to satisfy a version of the *no-signalling principle* [9], which states that the choice of measurement on one system cannot have a observable effect on the measurement results of a distinct system.

Finally, we note that any probabilistic theory can be viewed through "modal glasses", simply interpreting probabilities $p > 0$ as "possible" and $p = 0$ as "impossible". Thus, modal theories are generalizations of probabilistic theories. This generalization is actually an extension, since we will find modal theories that cannot be "resolved' to probabilistic ones. For situations that arise from systems in MQT, however, the situation is more complex. In the bipartite case we will show that a weak probabilistic resolution (which may assign $p = 0$ for a "possible" measurement result) can always be found.

2 Modal quantum theory

2.1 A modal world

The world of modal quantum theory is a world without probabilities. Probabilities are so familiar that it is worthwhile to consider more carefully what their absence entails.

In a probabilistic world, an event x is assigned a numerical probability $p(x)$ such that $0 \leq p(x) \leq 1$. The probabilities are normalized, so that

$$\sum_x p(x) = 1 \tag{1}$$

where the sum extends over a set of mutually exclusive and exhaustive events. Probabilities are related to statistical frequencies. Suppose we perform N independent trials of an experiment and observe event x to occur N_x times. Then with high probability,[2]

$$p(x) \approx \frac{N_x}{N}. \tag{2}$$

The possible results of an experiment may be labeled by numerical values v. The mean of such a random variable is given by

$$\langle v \rangle = \sum_v p(v)\, v. \tag{3}$$

In a *possibilistic* or *modal* world, we can only distinguish between possible and impossible events, but we do not assign any measure of likelihood to them. That is, we can identify a possible set

$$\mathcal{P} = \{x, x', \ldots\}. \tag{4}$$

The only "normalization" condition is the requirement that $\mathcal{P} \neq \emptyset$. If we perform an experiment many times, the set $\mathcal{R}$ of results that we see satisfies $R \subseteq P$. That is, every result we have actually seen is surely possible, but we can draw no definite conclusions about the possibility or impossibility of

---

[2]Note that the connection between probabilities and statistical frequencies is itself probabilistic! This highlights the difficulty in giving a non-circular operational interpretation of probability.

other results. Also, without any assignment of "weights" to the numerical results v of an experiment, we cannot compute a mean value $\langle v \rangle$.

The naive connection between probabilistic and modal pictures is that $x \in \mathcal{P}$ if and only if $p(x) \neq 0$. There are, however, some subtleties to be recognized. If we are assigning probabilities based on observed statistical frequencies, we cannot distinguish between a very rare event x (which may not have happened yet in our large but finite set of trials) and an impossible one. That is, we may be able to conclude that $p(x) \approx 0$ but not that x is impossible.

2.2 Basic axioms

The axioms for modal quantum theory are closely related to those of actual quantum theory, as we can see in Table 2.2. The axioms presented are for the most elementary versions of each theory; we will develop them further below. Even without this development, however, we can identify some interesting features of MQT. Consider, for instance, the case where $\mathcal{F}$ is a finite field. A system with finite-dimensional $\mathcal{V}$ has only a finite set of possible state vectors. There are only finitely many distinct measurements or time evolution maps for the system, and time evolution must proceed in discrete steps.

The simplest possible system in MQT is a "modal bit" or *mobit* [6], for which $\dim \mathcal{V} = 2$. If we also choose the smallest field $\mathcal{F} = \mathbb{Z}_2$, then there are just three non-zero vectors in $\mathcal{V}$, which we can denote $|0)$, $|1)$ and $|\sigma) = |0) + |1)$. The dual space $\mathcal{V}^*$ also has three vectors, so that

$$
\begin{array}{lll}
(a\,|0) = 1 & (a\,|1) = 0 & (a\,|\sigma) = 1 \\
(b\,|0) = 0 & (b\,|1) = 1 & (b\,|\sigma) = 1 \\
(c\,|0) = 1 & (c\,|1) = 1 & (c\,|\sigma) = 0
\end{array}
\qquad (5)
$$

Any pair of these dual vectors yields a basic measurement. There are thus three basic mobit measurements corresponding to the bases $X = \{(c|, (a|\}$, $Y = \{(b|, (c|\}$ and $Z = \{(a|, (b|\}$. The individual dual vectors in a measurement basis, which correspond to the results of the measurement, are called *effects*. It will sometimes be convenient to label the measurement results by $+$ and $-$, so we may write $(a| = (+_z| = (-_x|$, etc. If a mobit is in the state $|\sigma)$ and a Z-measurement is made, both outcomes $+_z$ and $-_z$ are possible.

As in AQT, we can compare MQT to a hypothetical hidden variable theory. Such a theory supposes that the system possesses some unknown

Actual quantum theory	Modal quantum theory
States. A system is described by a Hilbert space $\mathcal{H}$ over the field $\mathbb{C}$ of complex numbers. A state is a normalized $|\psi\rangle \in \mathcal{H}$.	**States.** A system is described by a vector space $\mathcal{V}$ over a field $\mathcal{F}$. A state is a non-zero $|\psi) \in \mathcal{V}$.
Measurements. A measurement is an orthonormal basis $\{|k\rangle\}$ for $\mathcal{H}$. Each basis element represents a measurement outcome. For state $|\psi\rangle$, the probability outcome k is $$p(k) = |\langle k\,|\psi\rangle|^2.$$	**Measurements.** A measurement is a basis $\{(k|\}$ for $\mathcal{V}^*$. Each basis element represents a measurement outcome. For state $|\psi)$, outcome k is possible if and only if $$(k\,|\psi) \neq 0.$$
Evolution. Over a given time interval, an isolated system evolves via a unitary operator U: $$|\psi\rangle \rightarrow U\,|\psi\rangle.$$	**Evolution.** Over a given time interval, an isolated system evolves via an invertible operator T: $$|\psi) \rightarrow T\,|\psi).$$
Composite systems. The state space for a composite system is the tensor product of subsystem spaces: $$\mathcal{H}^{(AB)} = \mathcal{H}^{(A)} \otimes \mathcal{H}^{(B)}.$$	**Composite systems.** The state space for a composite system is the tensor product of subsystem spaces: $$\mathcal{V}^{(AB)} = \mathcal{V}^{(A)} \otimes \mathcal{V}^{(B)}.$$

Table 1: Elementary axioms for AQT and MQT.

variable λ such that, for a given value of λ, the result of any measurement is determined. As in AQT, we cannot completely exclude all hidden variable theories, though we can show that some kinds of are inconsistent with MQT. For instance, consider a *non-contextual* hidden variable theory [10], in which (given λ) the question of whether a given effect $(e|$ will occur is independent of which other effects are present in the measurement basis. For a given value of λ, the theory would have to assign "yes" or "no" values to each of the dual vectors $(a|$, $(b|$ and $(c|$, such that any pair of them includes exactly one "yes". This is plainly impossible. We conclude that the pattern of possibilities for a mobit system in MQT cannot be reproduced by any non-contextual hidden variable theory.

This is essentially an MQT version of the famous Kochen-Specker theorem of AQT [10]. The MQT argument has a similar structure to the original (both can be cast as graph-coloring problems) but is radically simpler. Furthermore, the AQT version of the Kochen-Specker theorem only applies for $\dim \mathcal{H} \geq 3$, while the MQT version applies to any system of any dimension [11].

2.3 Entanglement

Composite systems in MQT may be in either product or entangled states. For instance, a pair of $\mathbb{Z}_2$-mobits has 15 possible states, of which 9 are product states and 6 are entangled. (For more complicated systems, the entangled states greatly outnumber the product states.)

Entangled states are marked by correlated measurement results. For example, consider the modal "singlet" state of two mobits:

$$|S) = |0, 1) - |1, 0) . \tag{6}$$

(The minus sign here allows us to generalize the state for any field $\mathcal{F}$. For $\mathcal{F} = \mathbb{Z}_2$, $-1 = +1$ and so $|S) = |0, 1) + |1, 0)$.) Note that, for any effect $(e|$,

$$(e, e | S) = (e | 0) (e | 1) - (e | 1) (e | 0) = 0. \tag{7}$$

Therefore, if we make the same measurement on both mobit subsystems, it is impossible that we obtain identical results.

The mobit measurements X, Y and Z yield nine possible joint measurements of a pair of mobits.[3] Let $(u, v | U, V)$ denote the situation in which

---

[3] There are also many measurements involving entangled effects.

measurements of U and V on two systems yield respective results u and v. Then we can summarize the measurement results for $|S\rangle$ as follows:

- If the same measurement is made on each mobit, the results must disagree. Thus $(+, +|X, X)$ is impossible, and so on.

- If different measurements are made on the two mobits, all but one of the joint results are possible. Thus, $(+, -|X, Y)$ is impossible, but $(+, +|X, Y)$, $(-, +|X, Y)$ and $(-, -|X, Y)$ are all possible.

For AQT, Bell showed that the correlations between entangled quantum systems were incompatible with any local hidden variable theory [2]. He did this by devising a statistical inequality that holds for any local hidden variable theory, but which can be violated by entangled quantum systems. Unfortunately, a similar approach based on probabilities and expectation values is not available in MQT.

Hardy devised an alternate argument for AQT based only on possibility and impossibility [3]. He constructed a non-maximally entangled state $|\Psi\rangle$ for a pair of qubits together with a set of measurements having the following properties:

- $(+, +|A, D)$ and $(+, +|B, C)$ are both impossible—that is, they have quantum probability $p = 0$.

- $(+, +|B, D)$ is possible $(p > 0)$.

- $(-, -|A, C)$ is impossible $(p = 0)$.

How might a local hidden variable theory account for this situation? Since $(+, +|B, D)$ is possible, we restrict our attention to the set H of hidden variable values that yield this result. The result of a measurement on one qubit is unaffected by the choice of measurement on the other (locality). Furthermore, no allowed values of the hidden variables can lead to $(+, +|A, D)$ or $(+, +|B, C)$. Thus, for values in H, we must obtain the results $(-, +|A, D)$ and $(+, -|B, C)$. These jointly imply that the result $(-, -|A, C)$ would be obtained for values in H. But this contradicts AQT, in which $(-, -|A, C)$ is impossible.

The very same argument applies to the state $|S\rangle$ in MQT, if we identify $A = X$, $B = Y$, $C = \bar{Z}$ (the negation of Z) and $D = \bar{Y}$. Thus we can conclude that no local hidden variable theory can account for the pattern of possible measurement outcomes generated by the entangled state $|S\rangle$.

However, this argument has a weakness, because it only applies to those situations in which the joint outcome $(+, +|B, D) = (+, -|Y, Y)$ actually occurs. In AQT, we can assign a finite probability $p > 0$ to this result, so we can confidently expect it to arise in a large enough sample. But in MQT, the statement that the joint result is possible does not allow us to draw any such conclusion. The MQT version of the Hardy argument therefore applies only to a situation that may not, in fact, ever occur.

A stronger argument may be constructed along the following lines [6]. We imagine that the MQT state $|S\rangle$ corresponds to a set H_S of possible values of a hidden variable. The variable controls the outcomes of possible measurements in a completely local way. For any particular value $h \in H_S$, the set of possible results of a measurement on one mobit depends only on the measurement choice for that mobit, not on the choice for the other mobit. Let $\mathcal{P}_h(E)$ be the set of possible results of a measurement of E for the hidden variable value h. Our locality assumption means that, given $V^{(1)}$ and $W^{(2)}$ measurements for the two mobits,

$$\mathcal{P}_h(V^{(A)}, W^{(B)}) = \mathcal{P}_h(V^{(A)}) \times \mathcal{P}_h(W^{(B)}), \tag{8}$$

the simple Cartesian product of separate sets $\mathcal{P}_h(V^{(1)})$ and $\mathcal{P}_h(W^{(2)})$. The MQT set of possible results arising from $|S\rangle$ should therefore be

$$\mathcal{P}(V^{(1)}, W^{(2)}|S) = \bigcup_{h \in H_S} \mathcal{P}_h(V^{(1)}) \times \mathcal{P}_h(W^{(2)}). \tag{9}$$

The individual sets $\mathcal{P}_h(V^{(1)})$, etc., are simultaneously defined for all of the measurements that can be made on either mobit. Therefore, we may consider the set

$$\begin{aligned}
\mathcal{J} &= \bigcup_{h \in H_S} \mathcal{P}_h(X^{(1)}) \times \mathcal{P}_h(Y^{(1)}) \times \mathcal{P}_h(Z^{(1)}) \\
&\quad \times \mathcal{P}_h(X^{(2)}) \times \mathcal{P}_h(Y^{(2)}) \times \mathcal{P}_h(Z^{(2)}). \tag{10}
\end{aligned}$$

There might be up to $2^6 = 64$ elements in $\mathcal{J}$. However, since $\mathcal{J}$ can only contain elements that agree with the properties of $|S\rangle$, we can eliminate many elements. For instance, the fact that corresponding measurements on the two mobits must give opposite results tells us that $(+, +, +, +, +, +)$ cannot be in $\mathcal{J}$, though $(+, +, +, -, -, -)$ might be. However, when we apply all of the properties of $|S\rangle$ in this way, we find the surprising result

10

that *all* of the elements of $\mathcal{J}$ are eliminated. *No* assignment of definite results to all six possible measurements can possibly agree with the correspondences obtained from the entangled MQT state $|S\rangle$. We therefore conclude that these correspondences are incompatible with any local hidden variable theory.

This argument can be recast in terms of a *pseudo-telepathy game* [12]. Two players, Alice and Bob, are separately asked questions drawn from a finite set. Their goal is to give answers that satisfy some joint criterion. The game is a pseudo-telepathy game if the goal could only be satisfied by classical players if they could communicate with each other. However, Alice and Bob may have a winning strategy if they share quantum entanglement. In our pseudo-telepathy game, Alice and Bob are each asked one of three questions (X, Y or Z), and their goal is to provide a joint answer consistent with the possible measurement outcomes of the entangled mobit state $|S\rangle$ described above. If Alice and Bob answer separately based on shared classical information, they cannot always win the game. If they share a mobit pair in $|S\rangle$, they can. (However, as we will see below in Section 5.3, this game has no perfect strategy in AQT.)

3 States and measurements

3.1 Generalized states and measurements in AQT

The axioms for MQT presented in Table 2.2 describe a "basic" version of the theory. In this section and the next, we will develop the theory to include more general kinds of states, measurements, and time evolution. Our development parallels the standard one in AQT [13], but also has many important differences.

In AQT, there are situations in which we cannot ascribe a definite state vector $|\psi\rangle$ to a system, either because its preparation is not completely known or because we have a subsystem of a larger composite system in an entangled state. In either case, we can construct a description of the situation from which we can make probabilistic predictions about the behavior of the system. We describe such *mixed states* by means of *density operators*.

Suppose, for instance that the system is prepared in one of several possible pure states, so that $|\psi_\alpha\rangle$ occurs with probability p_α. This mixture of states

is described by the density operator

$$\rho = \sum_\alpha p_\alpha \left| \psi_\alpha \right\rangle \! \left\langle \psi_\alpha \right|. \tag{11}$$

If we make a measurement on the system corresponding to an orthonormal basis $\{|k\rangle\}$, then the overall probability of the result k is

$$p(k) = \sum_\alpha p_\alpha \left| \langle k \, | \psi_\alpha \rangle \right|^2 = \langle k | \, \rho \, | k \rangle. \tag{12}$$

Thus, the density operator ρ is sufficient to predict the probability of any basic measurement result, given the probabilistic mixture of states.

Different mixtures of states can yield the same ρ, and thus yield the same statistical predictions. We therefore say that the different mixtures correspond to the same mixed state. Conversely, different density operators ρ and ρ' will lead to different statistical predictions for at least some measurements.

Density operators can also be used to describe a system that is part of a composite system. Given a joint state $|\Psi\rangle$ of RQ, we can construct a density operator for Q via the partial trace operation:

$$\rho = \mathrm{Tr}_{(\mathrm{R})} \left| \Psi \right\rangle \! \left\langle \Psi \right|. \tag{13}$$

Again, this density operator predicts the probabilities for any basic measurement on subsystem Q itself, according to the rule in Equation 12.

Every density operator arising from a mixture or a partial trace is a positive semidefinite operator of trace 1, and any such operator could arise in these ways. The set of positive semidefinite operators of trace 1 therefore constitutes our set of generalized states for a system.

We can also develop the concept of measurement in AQT. As a first step, we can "coarse-grain" a basic measurement, so that each outcome a corresponds to a projection operator Π_a (associated with the subspace of $\mathcal{H}$ spanned by the basis vectors $|k\rangle$ included in a). We can generalize further by supposing that we apply our measurement to a composite system and an *ancilla* system, which is regarded as part of the experimental apparatus. Then we find that each outcome a is associated with a positive semidefinite operator E_a, and that the probability of this outcome is

$$p(a) = \mathrm{Tr}\, \rho E_a. \tag{14}$$

The outcome operators E_a, sometimes called *effect operators*, sum to the identity:

$$\sum_a E_a = \mathbf{1}. \tag{15}$$

Our generalized model of measurement is thus a set $\{E_a\}$ of positive operators that satisfy 15. It can be further shown that any such set can be realized as a coarse-grained basic measurement on an extended system (i.e., they are feasible).

Finally, it is possible to give an axiomatic characterization of this development. The probability $p(a)$ of measurement result a is a functional of ρ, more completely written as $p(a) = p(a|\rho)$. The state ρ may arise as a mixture of two other states: $\rho = p_1\rho_1 + p_2\rho_2$. We infer that the probability of a for ρ should itself be a probabilistic combination:

$$p(a|\rho) = p_1 p(a|\rho_1) + p_2 p(a|\rho_2). \tag{16}$$

This motivates the axiom that $p(a|\rho)$ is a linear functional of ρ. Every such linear functional has the form of Equation 14 for some operator E_a. Since the probability of a must be real and non-negative for any state ρ, the operator E_a is positive semidefinite; and since the probabilities must always sum to 1, Equation 15 must also hold.

The developments of mixed states and generalized measurements in AQT thus illustrate the ideas of construction, feasibility and axiomatic characterization outlined in Subsection 1.2 above. We may regard this as a "complete" development within the theory. We are now ready to sketch the corresponding development in MQT.

3.2   Annihilators and mixed states

Modal quantum theory is based on a vector space $\mathcal{V}$ of states and its dual $\mathcal{V}^*$ containing effects. It is convenient to summarize here a few definitions and elementary results about the subspaces of $\mathcal{V}$ and $\mathcal{V}^*$ [14].

Subspaces of $\mathcal{V}$ form a lattice under the "meet" and "join" operations $\wedge$ and $\vee$, where $\mathsf{A} \wedge \mathsf{B} = \mathsf{A} \cap \mathsf{B}$ and $\mathsf{A} \vee \mathsf{B} = \langle \mathsf{A} \cup \mathsf{B} \rangle$. (Here $\langle X \rangle$ is the linear span of a set X.) The minimal subspace in this lattice is $\langle 0 \rangle$, the 0-dimensional null subspace of $\mathcal{V}$.

Given a set A of vectors in $\mathcal{V}$, the annihilator A° is the set of dual vectors in $\mathcal{V}^*$ that "annihilate" vectors in A. That is,

$$A^\circ = \{ (e| \in \mathcal{V}^* : (e|a) = 0 \text{ for all } |a) \in A \}. \tag{17}$$

This can be easily turned around to define the annihilator of a subset of the dual space $\mathcal{V}^*$. In this case, the annihilator would be a subset of $\mathcal{V}$.

Within modal quantum theory, if A is a set of states, then A° includes all effects that are impossible for every state in A. Dually, if A is a set of effects, then A° includes all states for which every one of the effects in A is impossible. In spaces of finite dimension, the annihilator of a set has several straightforward properties.

- The annihilator A° is a subspace.

- If $A \subseteq B$, then $B^\circ \subseteq A^\circ$.

- The set A and its span $\langle A \rangle$ have the same annihilator: $A^\circ = \langle A \rangle^\circ$.

- A and B have the same annihilator if and only if $\langle A \rangle = \langle B \rangle$.

- $(A \cup B)^\circ = A^\circ \wedge B^\circ$.

Finally, we note that the annihilator of the annihilator is a subspace of the original space, and in fact $A^{\circ\circ} = \langle A \rangle$. If A is a subspace, then $\mathsf{A}^{\circ\circ} = \mathsf{A}$.

3.3 Mixed states in MQT

In modal quantum theory, a pure state of a system is represented by a state vector $|\psi)$ in $\mathcal{V}$. How should we represent a mixed state? We approach this question first by considering mixtures of pure states. Since MQT does not involve probabilities, a mixture is merely a set of possible state vectors: $M = \{|\psi_1), |\psi_2), \ldots\}$. A particular measurement outcome is possible provided it is possible for at least one of the states in the mixture. That is, effect $(e|$ is possible provided it is *not* in the annihilator M°.

Two different mixtures M_1 and M_2 will thus predict exactly the same possible effects if and only if $M_1^\circ = M_2^\circ$, so that $\langle M_1 \rangle = \langle M_2 \rangle$. We say that two such mixtures yield the same *mixed state*, and we identify that state with the *subspace* $\mathsf{M} \subseteq \mathcal{V}$ spanned by the elements of the mixture.[4]

We can also consider mixtures of two or more mixed states. If M_1 and M_2 are two subspaces of $\mathcal{V}$ associated with two states, then $\mathsf{M}_1 \vee \mathsf{M}_2$ is the subspace associated with a mixture of the two. Since any non-null subspace

---

[4]This clarifies a point about pure states, that $|\psi)$ and $c|\psi)$ are operationally equivalent for any $c \neq 0$. The two vectors span the same one-dimensional subspace of $\mathcal{V}$.

of $\mathcal{V}$ can be written as the span of a set of state vectors, it is an allowed mixed state.

As in AQT, mixed states in MQT can also arise when a composite system is in an entangled pure state. Suppose that the composite system RQ is in a state $|\Psi^{(\mathrm{RQ})}\rangle$, and consider the joint effect $(r^{(\mathrm{R})}, q^{(\mathrm{Q})}| = (r^{(\mathrm{R})}| \otimes (q^{(\mathrm{Q})}|$, which is possible provided $(r^{(\mathrm{R})}, q^{(\mathrm{Q})}|\Psi^{(\mathrm{RQ})}) \neq 0$. We can make sense of this by defining $|\psi_r^{(\mathrm{Q})}\rangle = (r^{(\mathrm{R})}|\Psi^{(\mathrm{RQ})})$. That is, if we expand $|\Psi^{(\mathrm{RQ})}\rangle$ in a product basis $\{|a^{(\mathrm{R})}, b^{(\mathrm{Q})})\}$ we can write

$$|\psi_r^{(\mathrm{Q})}\rangle = (r^{(\mathrm{R})}| \left(\sum_{a,b} \Psi_{ab} |a^{(\mathrm{R})}, b^{(\mathrm{Q})}\rangle \right) = \sum_b \left(\sum_a \Psi_{ab} \, (r^{(\mathrm{R})}|a^{(\mathrm{R})}) \right) |b^{(\mathrm{Q})}\rangle. \quad (18)$$

(The vector $|\psi_r^{(\mathrm{RQ})}\rangle$ is independent of the choice of the $\{|a^{(\mathrm{R})}, b^{(\mathrm{Q})})\}$ basis chosen for this computation.) The joint effect $(r^{(\mathrm{R})}, q^{(\mathrm{Q})}|$ is possible provided $(q^{(\mathrm{Q})}|\psi_r^{(\mathrm{Q})}) \neq 0$. Thus, it makes sense to interpret $|\psi_r^{(\mathrm{Q})}\rangle$ as the conditional state of Q given the R-effect $(r^{(\mathrm{R})}|$ for the overall state $|\Psi^{(\mathrm{RQ})}\rangle$.[5]

To define the unconditional subsystem state for Q, we just define the Q-subspace of conditional states for all conceivable R-effects:

$$\mathsf{M}^{(\mathrm{Q})} = \{ (r^{(\mathrm{R})}|\Psi^{(\mathrm{RQ})}) : (r^{(\mathrm{R})}| \in \mathcal{V}^{(\mathrm{R})*} \}. \quad (19)$$

An R-measurement is a basis $\{(k^{(\mathrm{R})}|\}$ of R-effects. Since these span $\mathcal{V}^{(\mathrm{R})*}$, we can see that

$$\mathsf{M}^{(\mathrm{Q})} = \left\langle \{ |\psi_k^{(\mathrm{Q})}\rangle \} \right\rangle, \quad (20)$$

where $|\psi_k^{(\mathrm{Q})}\rangle = (k^{(\mathrm{R})}|\Psi^{(\mathrm{RQ})})$.

This has the following important consequence. Whatever measurement is made on subsystem R, the mixture of conditional Q-states is exactly $\mathsf{M}^{(\mathrm{Q})}$ from Equation 19. Thus, the choice of R-measurement by itself makes no observable difference in the observable properties of Q.

This analysis can be extended to the case where the composite system itself is in a mixed state. However, it is more convenient to delay that discussion until we have also generalized the concept of measurement in MQT.

---

[5]Of course, if $|\psi_r^{(\mathrm{Q})}\rangle = 0$, then it is not a legitimate state vector; but in this case, the R-effect $|r\rangle$ is impossible. The formal inclusion of such phantom conditional states makes no difference to our analysis.

3.4 Effects and measurements

To generalize effects and measurements in modal quantum theory, it is instructive to begin with an axiomatic characterization.

In the abstract, an effect is simply a map E that assigns each subspace M of $\mathcal{V}$ an element of $\{\text{possible}, \text{impossible}\}$. But suppose $\mathsf{M} = \mathsf{M}_1 \vee \mathsf{M}_2$, the mixture of states M_1 and M_2. Then $E(\mathsf{M})$ should be possible if E is possible for either M_1 or M_2. We therefore adopt this requirement as an axiom for any reasonable effect map E. To make this a consistent rule, we will have to adopt the sensible convention that $E(\langle 0 \rangle)$ is always impossible.

Our axiom is equivalent to the statement that $E(\mathsf{M})$ is impossible if and only if both $E(\mathsf{M}_1)$ and $E(\mathsf{M}_2)$ are impossible. Therefore, for a given E we can consider the subspace $\mathsf{Z}_E \subseteq \mathcal{V}$

$$\mathsf{Z}_E = \bigvee \{\mathsf{M} : E(\mathsf{M}) \text{ is impossible}\}. \tag{21}$$

We see that $E(\mathsf{M})$ is impossible if and only if M is a subspace of Z_E. The map E can therefore be completely characterized by the annihilator subspace $\mathsf{Z}_E^\circ \subseteq \mathcal{V}^*$.

A generalized effect in MQT is thus defined to be a subspace $\mathsf{E} \subseteq \mathcal{V}^*$. For a generalized mixed state M, we say that $\mathsf{E}(\mathsf{M})$ is impossible if $\mathsf{M} \subseteq \mathsf{E}^\circ$ and possible otherwise. That is,

$$\mathsf{E}(\mathsf{M}) = \begin{cases} \text{possible} & (e\,|m) \neq 0 & \text{for some } (e| \in \mathsf{E}, |m) \in \mathsf{M} \\ \text{impossible} & (e\,|m) = 0 & \text{for all } (e| \in \mathsf{E}, |m) \in \mathsf{M} \end{cases} \tag{22}$$

This subspace characterization of generalized effects in MQT is exactly what we expect from a constructive approach. Beginning with an ordinary measurement given by the basis $\{(k|\}$ for $\mathcal{V}^*$, we can construct a coarse-grained effect E from a subset of the $(k|$ dual vectors. This effect is associated with a subspace of $\mathcal{V}^*$ (the one spanned by the relevant basis vectors). Both axiomatic and constructive approaches yield the same mathematical representation for generalized effects.

A generalized measurement will be a collection $\{\mathsf{E}_a\}$ of generalized effects (subspaces of $\mathcal{V}^*$) associated with the potential results of the measurement process. Some result must always occur, so we impose the requirement that, for any state M, at least one effect must be possible—that is, M cannot lie in the annihilator of all the generalized effects. Thus,

$$\bigcap_a \mathsf{E}_a^\circ = \langle 0 \rangle, \tag{23}$$

and so the generalized effects must satisfy

$$\bigvee_a \mathsf{E}_a = \mathcal{V}^*. \tag{24}$$

This is our "normalization" condition for a generalized measurement in MQT.

In actual quantum theory, a theorem due to Neumark[15] guarantees that any generalized positive operator measurement on a system Q can be realized by a basic measurement on a larger system. It is not difficult to confirm that an exactly analogous result holds in MQT. Thus, our generalized measurements are all feasible, in the sense discussed in Section 1.2. The constructive and axiomatic approaches coincide, and so our development is once again "complete".

3.5 Conditional states

In AQT, any pure entangled state $|\Psi\rangle$ of system RQ can be written in a special form, the *Schmidt decomposition* [13], as follows:

$$|\Psi\rangle = \sum_k \sqrt{\lambda_k} \, |k^{(\mathrm{R})}\rangle \otimes |k^{(\mathrm{Q})}\rangle \tag{25}$$

where $\{|k^{(\mathrm{R})}\rangle\}$ and $\{|k^{(\mathrm{Q})}\rangle\}$ are orthonormal bases for the two systems. It is easy to see that these are the eigenbases for the subsystem states $\rho^{(\mathrm{R})}$ and $\rho^{(\mathrm{Q})}$, and that the coefficients λ_k are the eigenvalues. The Schmidt decomposition is unique except for degeneracy among the λ_k values and some choices of relative phases among the two bases.

The MQT analogue of the Schmidt decomposition can be found as follows. Suppose $|\Psi^{(\mathrm{RQ})})$ is a joint state for RQ with subsystem mixed states $\mathsf{M}^{(\mathrm{R})}$ and $\mathsf{M}^{(\mathrm{Q})}$. Let $d^{(\mathrm{R})} = \dim \mathsf{M}^{(\mathrm{R})}$ and $d^{(\mathrm{Q})} = \dim \mathsf{M}^{(\mathrm{Q})}$. We introduce an R-basis $\{|k^{(\mathrm{R})})\}$ for which the first $d^{(\mathrm{R})}$ elements form a basis for $\mathsf{M}^{(\mathrm{R})}$. This means we can write

$$|\Psi^{(\mathrm{RQ})}) = \sum_k |k^{(\mathrm{R})}) \otimes \left|\psi_k^{(\mathrm{Q})}\right), \tag{26}$$

where the sum only requires the first $d^{(\mathrm{R})}$ terms. The state of Q is thus $\mathsf{M}^{(\mathrm{Q})} = \left\langle\{\left|\psi_k^{(\mathrm{Q})}\right)\}\right\rangle$. Since $\mathsf{M}^{(\mathrm{Q})}$ is spanned by $d^{(\mathrm{R})}$ vectors, we conclude that $d^{(\mathrm{R})} \geq d^{(\mathrm{Q})}$. A symmetric argument establishes that $d^{(\mathrm{Q})} \geq d^{(\mathrm{R})}$, so the dimensions are equal. We therefore identify that $s = d^{(\mathrm{R})} = d^{(\mathrm{Q})}$ as the *Schmidt number* of the state $|\Psi^{(\mathrm{RQ})})$.

17

The s vectors $\{|\psi_k^{(\mathrm{Q})})\}$ span a space of dimension s, so they must be linearly independent. We can thus construct a Q-basis $\{|k)^{(\mathrm{Q})}\}$ in which $|k^{(\mathrm{Q})}) = |\psi_k^{(\mathrm{Q})})$ for $k \le s$. Then

$$|\Psi^{(\mathrm{RQ})}) = \sum_k |k^{(\mathrm{R})}) \otimes |k^{(\mathrm{Q})}), \qquad (27)$$

where the sum only includes s terms. This is a Schmidt decomposition for $|\Psi^{(\mathrm{RQ})})$. It is not unique, since we had the freedom to choose any basis for the mixed state (subspace) of one of the systems.

This has a useful consequence. Given a mixed state $\mathsf{M}^{(\mathrm{Q})}$ of Q, an entangled state $|\Psi^{(\mathrm{RQ})})$ of RQ that leads to this mixed state is called a *purification* of $\mathsf{M}^{(\mathrm{Q})}$ in RQ. Now consider two different purifications $|\Psi_1^{(\mathrm{RQ})})$ and $|\Psi_2^{(\mathrm{RQ})})$ for the same $\mathsf{M}^{(\mathrm{Q})}$. Fixing a common Q-basis $\{|k^{(\mathrm{Q})})\}$, we can write Schmidt decompositions for both purifications:

$$\left|\Psi_{1,2}^{(\mathrm{RQ})}\right) = \sum_k |k^{(\mathrm{R})}) \otimes \left|k_{1,2}^{(\mathrm{Q})}\right). \qquad (28)$$

The two R-bases are connected by an invertible operator on $\mathcal{V}^{(\mathrm{R})}$: $T\left|k_1^{(\mathrm{R})}\right) = \left|k_2^{(\mathrm{R})}\right)$. Thus, two purifications of $\mathsf{M}^{(\mathrm{Q})}$ in RQ are connected via

$$\left|\Psi_2^{(\mathrm{RQ})}\right) = (T^{(\mathrm{R})} \otimes \mathbf{1}^{(\mathrm{Q})}) \left|\Psi_1^{(\mathrm{RQ})}\right); \qquad (29)$$

that is, by an invertible transformation on R alone.

A theorem of Hughston, Jozsa and Wootters [16] (though earlier discussed by Schrödinger [17] and also by Jaynes [18]) relates mixtures to entangled states in AQT and characterizes those mixtures that can give rise to a given density operator ρ. An exactly analogous result holds in MQT. First, any mixture for $\mathsf{M}^{(\mathrm{Q})}$ can be realized as a mixture of conditional states arising from a purification of $\mathsf{M}^{(\mathrm{Q})}$. (The ability to realize different mixtures by a choice of measurement on the purifying subsystem is the MQT analogue of the familiar "steering" property of actual quantum theory [19].) Second, the elements of any two mixtures for a given mixed state are linear combinations of each other, with coefficients given by an invertible matrix of scalars. That is, if $\mathsf{M} = \langle\{|\psi_{k,1})\}\rangle = \langle\{|\psi_{k,2})\}\rangle$, then

$$|\psi_{l,2}) = \sum_k T_{lk} |\psi_{k,1}), \qquad (30)$$

18

where the T_{lk} form an invertible matrix.

We now return to the question of conditional states and subsystem states for composite systems in MQT. How do these ideas work out in the context of generalized states and effects?

Suppose the composite system RQ is in the joint state $\mathsf{M}^{(\mathrm{RQ})}$, and the effect subspace $\mathsf{E}^{(\mathrm{R})}$ is part of some measurement on R. The conditional state of Q given this effect, which we can denote $\mathsf{M}_E^{(\mathrm{Q})}$, is defined via a map $\mathsf{C}(\cdot|\cdot)$:

$$
\begin{aligned}
\mathsf{M}_E^{(\mathrm{Q})} &= \mathsf{C}(\mathsf{M}^{(\mathrm{RQ})}|\mathsf{E}^{(\mathrm{R})}) \\
&= \left\langle \left\{ (e^{(\mathrm{R})}|m^{(\mathrm{RQ})}) : (e^{(\mathrm{R})}| \in \mathsf{E}^{(\mathrm{R})}, |m^{(\mathrm{RQ})}) \in \mathsf{M}^{(\mathrm{RQ})} \right\} \right\rangle
\end{aligned}
\tag{31}
$$

If $\mathsf{M}_E^{(\mathrm{Q})} = \langle 0 \rangle$, then the effect $\mathsf{E}^{(\mathrm{R})}$ is impossible.

The map $\mathsf{C}(\cdot|\cdot)$ respects mixtures in both the joint state and the effect. That is,

$$
\mathsf{C}\big(\mathsf{M}_1^{(\mathrm{RQ})} \vee \mathsf{M}_2^{(\mathrm{RQ})}|\mathsf{E}^{(\mathrm{R})}\big) = \mathsf{C}\big(\mathsf{M}_1^{(\mathrm{RQ})}|\mathsf{E}^{(\mathrm{R})}\big) \vee \mathsf{C}\big(\mathsf{M}_2^{(\mathrm{RQ})}|\mathsf{E}^{(\mathrm{R})}\big) \tag{32}
$$

$$
\mathsf{C}\big(\mathsf{M}^{(\mathrm{RQ})}|\mathsf{E}_1^{(\mathrm{R})} \vee \mathsf{E}_2^{(\mathrm{R})}\big) = \mathsf{C}\big(\mathsf{M}^{(\mathrm{RQ})}|\mathsf{E}_1^{(\mathrm{R})}\big) \vee \mathsf{C}\big(\mathsf{M}^{(\mathrm{RQ})}|\mathsf{E}_2^{(\mathrm{R})}\big) \tag{33}
$$

Equation 31 generalizes the expression in Equation 18 for conditional states of a composite system. It can also be used to define the unconditional subsystem state $\mathsf{M}^{(\mathrm{Q})}$, which we denote like so:

$$
\mathsf{M}^{(\mathrm{Q})} = \mathsf{R}_{(\mathrm{R})}\big(\mathsf{M}^{(\mathrm{RQ})}\big) = \mathsf{C}\big(\mathsf{M}^{(\mathrm{RQ})}|\mathcal{V}^{(\mathrm{R})^*}\big). \tag{34}
$$

Since we take the linear span in Equation 31, we only need to consider spanning sets for $\mathsf{M}^{(\mathrm{RQ})}$ and $\mathsf{E}^{(\mathrm{R})}$. That is, if $\mathsf{M}^{(\mathrm{RQ})} = \langle\{|\mu^{(\mathrm{RQ})})\}\rangle$ and $\mathsf{E}^{(\mathrm{R})} = \langle\{(\eta^{(\mathrm{R})}|\}\rangle$, then $\mathsf{M}_E^{(\mathrm{Q})} = \langle\{(\eta^{(\mathrm{R})}|\mu^{(\mathrm{RQ})})\}\rangle$. This is useful in calculations.

4 Open system evolution

4.1 Type M maps

According to the evolution postulates given in Table 1, the time evolution of state vectors in either actual or modal quantum theory can be described by a linear operator—unitary in the case of AQT ($|\psi\rangle \to U|\psi\rangle$), invertible in the case of MQT ($|\phi) \to T|\phi)$). In either case it is straightforward to generalize

this to mixed states. The density operator in AQT evolves via $\rho \to U\rho U^\dagger$, and in MQT a subspace evolves according to

$$\mathsf{M} \to T\mathsf{M} = \{T \left| \phi \right\rangle : \left| \phi \right\rangle \in \mathsf{M}\}. \tag{35}$$

These postulates apply when the system in question is isolated. When a system is subject to noise or interaction with its environment, a more general description of time evolution is needed. In this section we trace this development.

Generalized operations can be of two types. *Conditional* operations do not take place with certainty but only happen when some objective condition (e.g., a measurement result) is observed. *Unconditional* operations are those that take place with certainty.

In actual quantum theory, a general operation is a map on density operators: $\rho \to \mathcal{E}(\rho)$. For an input density operator ρ, the output $\mathcal{E}(\rho)$ of an unconditional operation must also be a density operator—that is, a positive semidefinite operator of trace 1. For conditional operations, the output is subnormalized so that $p = \mathrm{Tr}\,\mathcal{E}(\rho)$ is the probability that the operation occurs. The map $\mathcal{E}$ must respect mixtures; that is,

$$\mathcal{E}\left(p_1\rho_1 + p_2\rho_2\right) = p_1\mathcal{E}(\rho_1) + p_2\mathcal{E}(\rho_2). \tag{36}$$

Thus $\mathcal{E}$ is a linear map on density operators. This is a powerful condition, since it allows us to extend $\mathcal{E}$ to a linear map on the space of all operators—that is, to a *superoperator*.

In MQT, a general operation $\mathcal{E}$ on a system is a map on the subspaces of $\mathcal{V}$: $\mathsf{M} \to \mathsf{M}' = \mathcal{E}(\mathsf{M})$. For an unconditional operation, the output of the map must always be a legitimate state, a non-null subspace. This means that $\mathcal{E}(\mathsf{M}) \neq \langle 0 \rangle$ for $\mathsf{M} \neq \langle 0 \rangle$. This requirement is relaxed for conditional operations. In that case, $\mathcal{E}(\mathsf{M}) = \langle 0 \rangle$ merely signifies that the condition of the operation cannot arise for the input state M.

General operations in MQT must also respect mixtures, meaning that

$$\mathcal{E}\left(\mathsf{M}_1 \vee \mathsf{M}_2\right) = \mathcal{E}(\mathsf{M}_1) \vee \mathcal{E}(\mathsf{M}_2). \tag{37}$$

(To maintain consistency, we adopt the convention that $\mathcal{E}(\langle 0 \rangle) = \langle 0 \rangle$.) The map $\mathcal{E}$ is not simply a linear superoperator, so we cannot easily extend it to inputs other than subspaces. However, Equation 37 is still an important requirement. We will call subspace maps that respect mixtures in this way *Type M* maps.

20

Throughout the rest of this section, we will only consider unconditional operations in both AQT and MQT. The generalization to conditional operations is not difficult and is left as an exercise.

4.2 Constructive approach

Consider a situation in which the system of interest S interacts with an external "environment" system E, where E is initially in some fixed state. In AQT, the dynamics of just such an *open* system S is described by a map $\mathcal{E}$ on density operators:

$$\rho \to \mathcal{E}(\rho) = \text{Tr}_{(E)} U \left(\rho \otimes |0\rangle\langle 0| \right) U^\dagger \tag{38}$$

where $|0\rangle$ is the initial standard state of E and U is a unitary operator on the composite system SE.

By analogy, in MQT the evolution of an open system that interacts with an environment (initial state $\mathsf{M}_0^{(E)}$) can be described by the map $\mathcal{E}^{(S)}$ such that

$$\mathcal{E}^{(S)}(\mathsf{M}^{(S)}) = \mathsf{R}_{(E)}(T^{(SE)}(\mathsf{M}^{(S)} \otimes \mathsf{M}_0^{(E)})) \tag{39}$$

where $\mathsf{R}_{(E)}$ is the subsystem state reduction defined by equation (34). Without loss of generality, we may suppose that $\mathsf{M}_0^{(E)}$ is one-dimensional, since any mixed environment state can have a purification in a larger environment. We refer to these maps defined by invertible linear evolution on a larger system as *Type I* maps.

In AQT there is a way of representing the map $\mathcal{E}$ without the explicit involvement of the environment E in Equation 38. Consider a particular basis $\{|e_k\rangle\}$ for the Hilbert space of the environment E. For each k define the operator A_k by

$$A_k |\phi\rangle = \langle e_k| U |\phi, 0\rangle \tag{40}$$

for any $|\phi\rangle$ in $\mathcal{H}^{(S)}$. Even though we have used the environment E and the interaction U in this definition, the A_k operators act on $\mathcal{H}^{(S)}$ alone. We may use the $\{|e_k\rangle\}$ basis to do the partial trace in Equation 38. Given a pure state input $|\phi\rangle$,

$$\begin{aligned} \mathcal{E}(|\phi\rangle\langle\phi|) &= \sum_k \langle e_k| U \left(|\phi\rangle\langle\phi| \otimes |0\rangle\langle 0| \right) U^\dagger |e_k\rangle \\ &= \sum_k A_k |\phi\rangle\langle\phi| A_k^\dagger. \end{aligned} \tag{41}$$

And in general,

$$\mathcal{E}(\rho) = \sum_k A_k \rho A_k^\dagger. \tag{42}$$

This is called an *operator-sum representation* or *Kraus representation* of the map $\mathcal{E}$, and the operators A_k are called *Kraus operators* [13].

For an unconditional operation, the Kraus operators satisfy a normalization condition. If $\rho' = \mathcal{E}(|\phi\rangle\langle\phi|)$ for a normalized pure state $|\phi\rangle$, then

$$
\begin{aligned}
\mathrm{Tr}\,\rho' &= \sum_k \langle\phi|\, A_k^\dagger A_k\, |\phi\rangle \\
&= \langle\phi|\left(\sum_k A_k^\dagger A_k\right)|\phi\rangle. \tag{43}
\end{aligned}
$$

Since $\mathrm{Tr}\,\rho' = 1$ for every normalized input state $|\phi\rangle$,

$$\sum_k A_k^\dagger A_k = \mathbf{1}. \tag{44}$$

We can describe the map $\mathcal{E}$ entirely in terms of Kraus operators. It is often much more convenient to describe $\mathcal{E}$ in this way without considering the actual environment E, which might be very large and complex. An operator-sum representation is a compact description of how E affects the evolution of the state of S.

We can make the analogous construction in MQT. The system S interacts with an environment E, initially in the state $\mathsf{M}_0^{(\mathrm{E})} = \langle|0^{(\mathrm{E})}\rangle\rangle$, via the invertible operator $T^{(\mathrm{SE})}$. Let $\{(e_k^{(\mathrm{E})}|\}$ be a basis for $\mathcal{V}^{(\mathrm{E})*}$ and define the operator A_k on $\mathcal{V}^{(\mathrm{S})}$ by

$$A_k\,|\phi^{(\mathrm{S})}) = \left(e_k^{(\mathrm{E})}\right|T^{(\mathrm{SE})}\,|\phi^{(\mathrm{S})}, 0^{(\mathrm{E})}). \tag{45}$$

Given an initial S-state $\mathsf{M}^{(\mathrm{S})}$ spanned by state vectors $|m^{(\mathrm{S})})$, we have

$$
\begin{aligned}
\mathcal{E}^{(\mathrm{S})}\left(\mathsf{M}^{(\mathrm{S})}\right) &= \mathsf{R}_{(\mathrm{E})}(T^{(\mathrm{SE})}(\mathsf{M}^{(\mathrm{S})} \otimes \mathsf{M}_0^{(\mathrm{E})})) \\
&= \left\langle\left(e_k^{(\mathrm{E})}\right|T^{(\mathrm{SE})}\,|m^{(\mathrm{S})}, 0^{(\mathrm{E})})\right\rangle \\
&= \langle A_k\,|m^{(\mathrm{S})})\rangle \\
\mathcal{E}^{(\mathrm{S})}\left(\mathsf{M}^{(\mathrm{S})}\right) &= \bigvee_k A_k \mathsf{M}^{(\mathrm{E})}. \tag{46}
\end{aligned}
$$

The output of $\mathcal{E}^{(\mathrm{S})}$ acting on $\mathsf{M}^{(\mathrm{S})}$ is a mixture of images of $\mathsf{M}^{(\mathrm{S})}$ under the linear operators A_k. Equation 46 is the MQT analogue of the Kraus representation in Equation 42. If a map $\mathcal{E}^{(\mathrm{S})}$ has a representation of this type, we

22

say that it is a *Type L* map. (Note that we have shown that all Type I maps are also Type L.)

The individual operators A_k are not necessarily invertible. However, if $\mathcal{E}$ represents an unconditional operation on the MQT system S, then any non-null subspace M must evolve to a non-null subspace $\mathcal{E}(\mathsf{M})$. Thus, the A_k operators must satisfy

$$\bigcap_k \ker A_k = \langle 0 \rangle, \tag{47}$$

the MQT analogue of the normalization condition in Equation 44.

4.3 Axiomatic characterization

In AQT, every "physically reasonable" dynamical evolution map for an open system has both a unitary and a Kraus representation. Similarly, every "physically reasonable" evolution map in MQT is both Type I and Type L. To make sense of this claim, we must explain what is meant by a "physically reasonable" map.

Let us begin by reviewing the argument in AQT. A physically reasonable map $\mathcal{E}$ must be linear in the input density operator, making it a superoperator (an element of $\mathcal{B}(\mathcal{B}(\mathcal{H}))$). Furthermore, the output of $\mathcal{E}$ must be a valid density operator for any valid input state. This immediately implies two properties:

- $\mathcal{E}$ must be a *positive* map, in the sense that it maps positive operators to positive operators.

- $\mathcal{E}$ must be a *trace-preserving* map, so that $\operatorname{Tr} \mathcal{E}(A) = \operatorname{Tr} A$ for all operators A.

(Both properties are easy to prove in general, since any operator can be written as a linear combination of density operators.)

These two conditions are not sufficient to characterize "physically reasonable" linear maps, because there are positive, trace-preserving maps that cannot correspond to the time evolution of an quantum system. The easiest example arises for a simple qubit system. The Pauli operators X, Y and Z, together with the identity $\mathbf{1}$, form an operator basis. The following map $\mathcal{T}$ is positive:

$$\begin{aligned} \mathcal{T}(\mathbf{1}) &= \mathbf{1} & \mathcal{T}(X) &= X \\ \mathcal{T}(Y) &= Y & \mathcal{T}(Z) &= -Z \end{aligned}. \tag{48}$$

However, $\mathcal{T}$ does not describe the possible evolution of the state of an open qubit system. The reason is that the qubit is not necessarily alone in the universe. We may consider a second independent qubit whose state evolves according to the identity map $\mathcal{I}$. The composite system evolves according to the map $\mathcal{T} \otimes \mathcal{I}$. However, this extended map is *not* positive: entangled input states may map to operators having some negative eigenvalues. (See [13] for details.)

We need stronger property to characterize "physically reasonable" evolution maps for an open quantum system. The structure of the example just described provides a clue to what this stronger property looks like.

The map $\mathcal{E}$ for a system is said to be *completely positive* if $\mathcal{E} \otimes \mathcal{I}$ is positive whenever we append an independent quantum system. This means that, for any initial pure state $|\Psi\rangle$ of the composite system, the operator $\mathcal{E} \otimes \mathcal{I}(|\Psi\rangle\langle\Psi|)$ is positive. Since any system may be part of a composite system, we require that every "physically reasonable" map describing open system evolution must be linear, trace-preserving and completely positive. The importance of this requirement is shown by the following theorem.

> **Representation theorem for generalized dynamics in AQT.**
> Let Q be a quantum system and $\mathcal{E}$ be a map on Q-operators. The following conditions are equivalent.
>
> **(a)** $\mathcal{E}$ is a linear, trace-preserving, completely positive map.
>
> **(b)** $\mathcal{E}$ has a "unitary representation". That is, we can introduce an environment system E, an initial environment state $|0\rangle$ and a joint unitary evolution U on QE so that
>
> $$\mathcal{E}(G) = \text{Tr}_{(E)} U \left(G \otimes |0\rangle\langle 0| \right) U^\dagger. \tag{49}$$
>
> **(c)** $\mathcal{E}$ has a Kraus representation. That is, we can find operators A_k such that
> $$\mathcal{E}(G) = \sum_k A_k G A_k^\dagger. \tag{50}$$
>
> The Kraus operators satisfy the normalization condition of Equation 44.

Once again, the constructive approach (unitary dynamics on a larger system) and the axiomatic characterization (linear, trace-preserving, completely positive maps) lead us to the same generalized unconditional operations in AQT.

The representation theorem is a powerful and fundamental result in AQT. Details of its proof are given in Appendix D of [13].

So much for AQT. Can we find an analogous axiomatic characterization for "reasonable" state evolution in MQT? This is a tricky question, and not just because we lack access to an actual MQT world. A proposed evolution map $\mathcal{E}$ will map subspaces to subspaces, rather than operators to operators. Furthermore, the underlying field $\mathcal{F}$ may not include the notion of "positive elements". In a field of non-zero characteristic, any element added to itself sufficiently many times yields 0. Thus, in MQT there may be no analogue to the notion of a "positive map".

Nevertheless, there is a close analogue to the property of complete positivity that does make sense in MQT. As in AQT, this condition governs how a map $\mathcal{E}$ extends to one that applies to a larger composite system. Briefly, we require that an extension exists that commutes with the conditioning operation described in Subsection 3.5. Here is a more precise definition: We say that the subspace map $\mathcal{E}^{(\mathrm{S})}$ for modal quantum system S is *Type E* if for any other system R there exists a joint subspace map $\mathcal{E}^{(\mathrm{RS})}$ such that

$$\mathcal{E}^{(\mathrm{S})}(\mathsf{C}(\mathsf{M}^{(\mathrm{RS})}|\mathsf{E}^{(\mathrm{R})})) = \mathsf{C}(\mathcal{E}^{(\mathrm{RS})}(\mathsf{M}^{(\mathrm{RS})})|\mathsf{E}^{(\mathrm{R})}) \tag{51}$$

for any RQ-state $\mathsf{M}^{(\mathrm{RQ})}$ and R-effect $\mathsf{E}^{(\mathrm{R})}$. In other words, for a Type E map $\mathcal{E}^{(\mathrm{S})}$ and a system R, we can find a map $\mathcal{E}^{(\mathrm{RS})}$ so that the following diagram always commutes:

$$
\begin{array}{ccc}
\mathsf{M}^{(\mathrm{RS})} & \xrightarrow{\ \mathsf{C}(\cdot|\mathsf{E}^{(\mathrm{R})})\ } & \mathsf{M}^{(\mathrm{S})} \\[2mm]
{\scriptstyle \mathcal{E}^{(\mathrm{RS})}}\Big\downarrow & & \Big\downarrow{\scriptstyle \mathcal{E}^{(\mathrm{S})}} \\[2mm]
\mathcal{E}^{(\mathrm{RS})}(\mathsf{M}^{(\mathrm{RS})}) & \xrightarrow{\ \mathsf{C}(\cdot|\mathsf{E}^{(\mathrm{R})})\ } & \mathcal{E}^{(\mathrm{S})}(\mathsf{M}^{(\mathrm{S})})
\end{array}
\tag{52}
$$

We will require that any "reasonable" state evolution in MQT must be Type E. What is the motivation for such a condition? Suppose the state of system S evolves according to $\mathcal{E}^{(\mathrm{S})}$. It is always reasonable to suppose that another system R exists in the MQT universe. We imagine that R and S can be "independent" of one another—they might, for instance, be very far apart in space. The joint system RS is initially in the state $\mathsf{M}^{(\mathrm{RS})}$. The evolution of RS is described by some joint map $\mathcal{E}^{(\mathrm{RS})}$ that reflects the independence of the subsystems. We now imagine two experimental procedures.

25

- A measurement is made on system R, with the objective result corresponding to effect $\mathsf{E}^{(\mathrm{R})}$. Under this condition, S is in the state $\mathsf{C}(\mathsf{M}^{(\mathrm{RS})}|\mathsf{E}^{(\mathrm{R})})$. Now the dynamical evolution acts, so that the final state of S is $\mathcal{E}^{(\mathrm{S})}\left(\mathsf{C}(\mathsf{M}^{(\mathrm{RS})}|\mathsf{E}^{(\mathrm{R})})\right)$.

- The dynamical evolution acts, leading to the joint final state $\mathcal{E}^{(\mathrm{RS})}(\mathsf{M}^{(\mathrm{RS})})$. Now the measurement is made on system R, with the objective result corresponding to the effect $\mathsf{E}^{(\mathrm{R})}$. The conditional state of S is $\mathsf{C}(\mathcal{E}^{(\mathrm{RS})}(\mathsf{M}^{(\mathrm{RS})})|\mathsf{E}^{(\mathrm{R})})$.

Intuitively, if R and S are completely independent (and perhaps widely separated), the final S state should be independent of whether the measurement on R is performed before or after the evolution of S. This is exactly the requirement for a Type E map.

4.4 Representation theorem for MQT

We now prove a result analogous to the representation theorem for AQT. Formally, we will show that

> **Representation theorem for generalized dynamics in MQT.**
> Let S be a modal quantum system and $\mathcal{E}$ be a Type M map on subspaces for F. Then the following conditions are equivalent.
>
> **(a)** $\mathcal{E}$ is Type E; that is, it can be extended in a way that commutes with the conditional operation.
>
> **(b)** $\mathcal{E}$ is Type I; that is, it can be expressed as invertible linear evolution on a larger system.
>
> **(c)** $\mathcal{E}$ is Type L; that is, it can be expressed as a mixture of linear maps satisfying Equation 47.

As with the AQT result, this theorem is a strong characterization of the "reasonable" evolution maps in modal quantum theory. We have argued that all reasonable maps are Type E, and any Type I map is realizable by familiar linear evolution. Constructive and axiomatic approaches coincide.

We will prove the equivalence of the three conditions by establishing the cyclic implication L $\Rightarrow$ I $\Rightarrow$ E $\Rightarrow$ L. The first two implications are straightforward; the last requires a bit more work. Throughout, we take $\mathcal{E}^{(\mathrm{S})}$ to be a Type M map on S.

L ⇒ I: First, assume that $\mathcal{E}^{(\mathrm{S})}$ is Type L. This means that there is a set of linear operators $\{A_k\}$ that yield $\mathcal{E}^{(\mathrm{S})}$ according to Equation 46, and that these operators satisfy the normalization requirement (Equation 47). Now we introduce an environment system E whose dimension is equal to the number of A_k operators. We fix an initial E-state $|0^{(\mathrm{E})})$ and a basis $\{|k^{(\mathrm{E})})\}$.

The set of SE states of the form $|\phi^{(\mathrm{S})}, 0^{(\mathrm{E})})$ constitute a subspace. Define the operator $T^{(\mathrm{SE})}$ on this subspace by

$$T^{(\mathrm{SE})}|\phi^{(\mathrm{S})}, 0^{(\mathrm{E})}) = \sum_k A_k |\phi^{(\mathrm{E})}) \otimes |k^{(\mathrm{E})}). \tag{53}$$

Because of the normalization requirement on the A_k operators, the right-hand side is never zero. Thus, the operator $T^{(\mathrm{SE})}$ is one-to-one on the subspace, and so we may extend it to an invertible operator on the whole of $\mathcal{V}^{(\mathrm{SE})}$. Given a mixed state $\mathsf{M}^{(\mathrm{S})}$, it is straightforward to show that

$$\mathsf{R}_{(\mathrm{E})}\left(T^{(\mathrm{SE})}(\mathsf{M}^{(\mathrm{S})} \otimes \mathsf{M}_0^{(\mathrm{E})})\right) = \langle\{A_k\mathsf{M}^{(\mathrm{S})}\}\rangle = \mathcal{E}^{(\mathrm{S})}(\mathsf{M}^{(\mathrm{S})}), \tag{54}$$

where $\mathsf{M}_0^{(\mathrm{E})} = \langle|0^{(\mathrm{E})})\rangle$. The map $\mathcal{E}^{(\mathrm{S})}$ is therefore Type I.

I ⇒ E: Now assume that $\mathcal{E}^{(\mathrm{S})}$ is Type I, so that it is given by invertible evolution on the extended system SE as above. For any additional system R we define the map

$$\mathcal{E}^{(\mathrm{RS})}(\mathsf{M}^{(\mathrm{RS})}) = \mathsf{R}_{(\mathrm{E})}\left[(1^{(\mathrm{R})} \otimes T^{(\mathrm{SE})})(\mathsf{M}^{(\mathrm{RS})} \otimes \mathsf{M}_0^{(\mathrm{E})})\right]. \tag{55}$$

As we have already remarked, the reduction operation $\mathsf{R}_{(\mathrm{E})}$ is an "unconditional" conditioning operation. A direct application of the definition in Equation 31 shows that iterated reduction with respect to independent subsystems (in our case, R and E) can be done in any order. This establishes that every Type I map must also be Type E.

E ⇒ L: It only remains to prove that Type E implies Type L. Let $\mathcal{E}^{(\mathrm{S})}$ be a Type E map for a modal quantum system S, which is represented by a vector space of finite dimension $\dim \mathcal{V}^{(\mathrm{S})} = d$. We append an identical quantum system R and consider the maximally entangled state

$$|\Phi^{(\mathrm{RS})}) = \sum_k |k^{(\mathrm{R})}, k^{(\mathrm{S})}). \tag{56}$$

Any initial state $|\psi^{(\mathrm{S})})$ of S could arise in the following way. The system RS is initially in the entangled state $|\Phi^{(\mathrm{RS})})$, and then a measurement is performed

in R. The resulting state of S, conditional on the particular measurement outcome for R, happens to be $|\psi^{(\mathrm{S})})$.

We can do this more explicitly. Given $|\psi^{(\mathrm{S})}) = \sum_k g_k |k^{(\mathrm{S})})$, we can construct the R-effect

$$\left(\tilde{\psi}^{(\mathrm{R})}\right| = \sum_k g_k \left(k^{(\mathrm{R})}\right|. \tag{57}$$

If $\left(\tilde{\psi}^{(\mathrm{R})}\right|$ corresponds to one outcome of a basic measurement on R, then the associated conditional state of Q is $|\psi^{(\mathrm{Q})})$.

This reasoning can be generalized to mixed states. First, note that $\mathcal{M}^{(\mathrm{RS})} = \langle|\Phi^{(\mathrm{RS})})\rangle$ is the one-dimensional mixed state that corresponds to the "fully entangled" state $|\Phi^{(\mathrm{RS})})$. Now consider a general mixed Q-state

$$\mathsf{G}^{(\mathrm{Q})} = \left\langle \left\{ |g^{(\mathrm{Q})}) = \sum_k g_k |k^{(\mathrm{S})}) : (g_k) \in G \right\} \right\rangle, \tag{58}$$

where G is a set of d-tuples (g_k) of elements of $\mathcal{F}$. Now define the R-effect

$$\Gamma^{(\mathrm{R})} = \left\langle \left\{ (g^{(\mathrm{R})}| = \sum_k g_k \left(k^{(\mathrm{R})}\right| : (g_k) \in G \right\} \right\rangle. \tag{59}$$

Then

$$\mathsf{G}^{(\mathrm{Q})} = \mathsf{C}(\mathcal{M}^{(\mathrm{RS})}|\Gamma^{(\mathrm{R})}), \tag{60}$$

since $|g^{(\mathrm{Q})}) = (g^{(\mathrm{R})}|\Phi^{(\mathrm{RS})})$.

We use this machinery to construct a Type L representation the Type E map $\mathcal{E}^{(\mathrm{S})}$. Let $\mathcal{E}^{(\mathrm{RS})}$ be an extension of $\mathcal{E}^{(\mathrm{S})}$. Let $\left\{|m_\lambda^{(\mathrm{RS})})\right\}$ be a set of RS states such that $\mathcal{E}^{(\mathrm{RS})}(\mathcal{M}^{(\mathrm{RS})}) = \langle\{|m_\lambda^{(\mathrm{RS})})\}\rangle$ (where λ runs over some index set Λ). Given states $|g^{(\mathrm{S})})$ and associated effects $(g^{(\mathrm{R})}|$ as described above, we define A_λ as follows:

$$A_\lambda^{(\mathrm{S})} |g^{(\mathrm{S})}) = \left(g^{(\mathrm{R})}\middle|m_\lambda^{(\mathrm{RS})}\right). \tag{61}$$

We now have

$$\begin{aligned}
\mathcal{E}^{(\mathrm{S})}(\mathsf{G}^{(\mathrm{S})}) &= \mathcal{E}^{(\mathrm{S})}(\mathsf{C}(\mathcal{M}^{(\mathrm{RS})}|\Gamma^{(\mathrm{R})})) \\
&= \mathsf{C}(\mathcal{E}^{(\mathrm{RS})}(\mathcal{M}^{(\mathrm{RS})})|\Gamma^{(\mathrm{R})}) \\
&= \mathsf{C}(\langle\{m_\lambda^{(\mathrm{RS})}\}\rangle|\Gamma^{(\mathrm{R})}) \\
&= \langle\{(g^{(\mathrm{R})}|m_\lambda^{(\mathrm{RS})})\}\rangle \\
&= \langle\{A_\lambda^{(\mathrm{S})}|g^{(\mathrm{S})})\}\rangle \\
\mathcal{E}^{(\mathrm{S})}(\mathsf{G}^{(\mathrm{S})}) &= \bigvee_\lambda A_\lambda^{(\mathrm{S})}(\mathsf{G}^{(\mathrm{S})}).
\end{aligned} \tag{62}$$

Thus, any Type E map is also Type L.

We see that Type E maps in MQT play a role parallel to CP maps in actual quantum theory. In fact, the connection is stronger than this. We can adapt the definition of Type E maps to AQT: A linear map $\mathcal{E}^{(Q)}$ on density operators for Q is Type E provided there exists an extended map $\mathcal{E}^{(RQ)}$ on density operators of RQ that commutes with the formation of conditional Q states from effects on R. It is not hard to show that this condition is *equivalent* to complete positivity of $\mathcal{E}^{(Q)}$ and thus implies the existence of unitary ("Type L") and Kraus ("Type L") representations in AQT.

With the (now finished) proof of the representation theorem, we have completed our development of modal quantum theory to include generalized states, measurements, and dynamical evolution. This development has included both constructive approaches (based on the basic axioms in Table 2.2) and axiomatic characterizations. The two routes lead to the same place, a fact that gives us confidence that we have arrived at a "complete" development of the theory. Further generalization will necessarily involve an extension of MQT to a more general type of theory.

5   Generalized modal theories

5.1   Possibility tables for two systems

In the study of the conceptual foundations of actual quantum theory, it is useful to consider AQT as an example of a more general class of probabilistic theories. Consider a system comprising two subsystems, designated 1 and 2. We can choose to make any of several possible measurements on each system and obtain various joint results with various probabilities. That is, our theory allows us to compute probabilities of the form $p(x, y | X^{(1)}, Y^{(2)})$, the probability of the joint outcome (x, y) given the choice of measurement $X^{(1)}$ on system 1 and $Y^{(2)}$ on system 2.

The state of the composite system can thus be described by a collection of joint probability distributions. These may be organized as a table. The rows and columns of the table correspond to the possible measurements on

systems 1 and 2, respectively, like so:

$$(63)$$

The theory is characterized by the set of possible states—that is, the possible collections of distributions in the table.[6]

All of the tables we consider satisfy the *no-signalling principle* which can be stated as follows [9]. For any choice of measurements $A^{(1)}$, $B^{(1)}$, $C^{(2)}$ and $D^{(2)}$,

$$
\begin{aligned}
p\left(a|A^{(1)}\right) &= \sum_c p\left(a,c|A^{(1)},C^{(2)}\right) &= \sum_d p\left(a,d|A^{(1)},D^{(2)}\right) \\
p\left(c|C^{(1)}\right) &= \sum_a p\left(a,c|A^{(1)},C^{(2)}\right) &= \sum_b p\left(b,c|B^{(1)},C^{(2)}\right).
\end{aligned}
\tag{64}
$$

That is, the choice of system 2 measurement does not affect the overall probability of a system 1 outcome, and *vice versa*. In the table of joint distributions in Equation 63, this means that any two distributions in the same row are connected, in that their sub-rows sum to the same values. A similar connection exists within each column as well.

Collections of distributions arising from a composite system in AQT satisfy the no-signalling principle. However, there are tables satisfying this principle that could not arise in AQT. These include examples of "states" of a composite system that are "more entangled" than quantum mechanics allows [20].

We can adapt this approach to construct generalized modal theories that extend MQT. The state of a composite system in a general modal theory would be a table similar to the one in Equation 63, except that the the

---

[6]AQT also allows "entangled" measurements on composite systems, measurements which cannot be reduced to separate measurements on the subsystems. Probabilities for non-entangled measurements, however, are sufficient to characterize the joint state of the system, so we restrict our attention to those.

individual "distributions" only indicate which joint outcomes are possible. We use the symbol X to denote a possible outcome, and a blank space for an impossible outcome. For instance, the table for a pair of mobits in $\mathbb{Z}_2$-MQT has just three rows and columns, corresponding to the three possible basic measurements for each mobit. For the entangled modal state $|S) = |0, 1) - |1, 0)$, we have

$$
\mathcal{S} =
\begin{bmatrix}
 & X^{(2)} & Y^{(2)} & Z^{(2)} \\[4pt]
X^{(1)} & \begin{array}{|c|c|}\hline \ & X \\\hline X & \ \\\hline\end{array} & \begin{array}{|c|c|}\hline X & \ \\\hline X & X \\\hline\end{array} & \begin{array}{|c|c|}\hline X & X \\\hline \ & X \\\hline\end{array} \\[10pt]
Y^{(1)} & \begin{array}{|c|c|}\hline X & X \\\hline \ & X \\\hline\end{array} & \begin{array}{|c|c|}\hline \ & X \\\hline X & \ \\\hline\end{array} & \begin{array}{|c|c|}\hline X & \ \\\hline X & X \\\hline\end{array} \\[10pt]
Z^{(1)} & \begin{array}{|c|c|}\hline X & \ \\\hline X & X \\\hline\end{array} & \begin{array}{|c|c|}\hline X & X \\\hline \ & X \\\hline\end{array} & \begin{array}{|c|c|}\hline \ & X \\\hline X & \ \\\hline\end{array}
\end{bmatrix}
\tag{65}
$$

This table, like all tables arising from MQT systems, satisfies a modal version of the no-signalling principle. The question of whether a particular subsystem result is possible does not depend on what measurement is chosen for the other subsystem. Thus, if an X occurs in the table, at least one X must occur in the corresponding sub-rows to the right and left, and in the corresponding sub-columns above and below. We will only consider general modal theories satisfying the modal no-signalling principle.

In a general probabilistic theory, we can take the convex combination of two states and derive a "mixed" state. The set of allowed states is therefore a convex set. In a general modal theory, the mixture of two tables $\mathcal{R}$ and $\mathcal{T}$ is simply $\mathcal{R} \vee \mathcal{T}$, the table in which a joint outcome is possible if it is possible in either $\mathcal{R}$ or $\mathcal{T}$. (This corresponds to the usual mixture of states in MQT.) Finally, there is a natural partial ordering on states in a general modal theory. We say that $\mathcal{R} \preceq \mathcal{T}$ provided every possible result in $\mathcal{R}$ is also possible in $\mathcal{T}$.

5.2 Popescu-Rohrlich boxes

Every generalized probabilistic table can be converted into a generalized modal table by replacing non-zero probabilities with X and zero probabil-

ities with blanks. A table obeying the probabilistic no-signalling principle automatically yields one that obeys the modal version.

We use this idea to create a modal version of an important example of a generalized probabilistic model, the "nonlocal box" proposed by Popescu and Rohrlich [20]. This *PR box* satisfies the no-signalling principle but is in a sense more entangled than allowed by AQT. The modal version $\mathcal{P}$ looks like this:

$$
\mathcal{P} =
\begin{array}{c|cc}
 & C^{(2)} & D^{(2)} \\
A^{(1)} & \begin{array}{|c|c|}\hline \mathsf{X} & \\\hline & \mathsf{X}\\\hline\end{array} & \begin{array}{|c|c|}\hline \mathsf{X} & \\\hline & \mathsf{X}\\\hline\end{array} \\
B^{(1)} & \begin{array}{|c|c|}\hline \mathsf{X} & \\\hline & \mathsf{X}\\\hline\end{array} & \begin{array}{|c|c|}\hline & \mathsf{X}\\\hline \mathsf{X} & \\\hline\end{array}
\end{array}
\tag{66}
$$

We can summarize this pattern of possibilities in a simple way: For the measurement combinations (A, C), (A, D) and (B, C) the joint measurement results must always agree, but for (B, D) they always disagree. (The probabilistic PR box replaces X with probability $1/2$ in Equation 66.)

The PR box table $\mathcal{P}$ is minimal. That is, if any table $\mathcal{R}$ of similar dimensions satisfies the no-signalling principle, and if $\mathcal{R} \preceq \mathcal{P}$, then $\mathcal{R} = \mathcal{P}$.

Could the PR box table $\mathcal{P}$ in Equation 66 arise from a composite system described by MQT? In fact, it cannot. Since $\mathcal{P}$ is minimal, it suffices to consider only pure states for system 12 together with measurements having non-overlapping effects. That is, the measurement $A^{(1)}$ consists of two effects (subspaces of $\mathcal{V}^{(1)*}$) A_+ and A_- such that $\mathsf{A}_+ \cap \mathsf{A}_- = \langle 0 \rangle$, and so on.

Suppose $|\Psi)$ is a modal quantum state that leads to the PR box table $\mathcal{P}$ in Equation 66. As shown in the Appendix, the upper-left quarter of the table tells us that

$$|\Psi) = |\Psi_+) + |\Psi_-), \tag{67}$$

where these two non-zero parts of $|\Psi)$ satisfy

$$|\Psi_+) \in \mathsf{A}_-^\circ \otimes \mathsf{C}_-^\circ \quad \text{and} \quad |\Psi_-) \in \mathsf{A}_+^\circ \otimes \mathsf{C}_+^\circ. \tag{68}$$

The same state vector $|\Psi)$ gives rise to the possibilities in the upper-right quarter of $\mathcal{P}$ also. From this we can conclude that

$$|\Psi_+) \in \mathsf{A}_-^\circ \otimes \mathsf{D}_-^\circ \quad \text{and} \quad |\Psi_-) \in \mathsf{A}_+^\circ \otimes \mathsf{D}_+^\circ. \tag{69}$$

We can continue around the table $\mathcal{P}$, arriving at the following facts:

$$|\Psi_+) \in \mathsf{B}_+^\circ \otimes \mathsf{D}_-^\circ \quad \text{and} \quad |\Psi_-) \in \mathsf{B}_-^\circ \otimes \mathsf{D}_+^\circ. \tag{70}$$

$$|\Psi_+) \in \mathsf{B}_+^\circ \otimes \mathsf{C}_+^\circ \quad \text{and} \quad |\Psi_-) \in \mathsf{B}_-^\circ \otimes \mathsf{C}_-^\circ. \tag{71}$$

This last pair of statements allows us to return to the upper-left corner, concluding that

$$|\Psi_+) \in \mathsf{A}_+^\circ \otimes \mathsf{C}_+^\circ \quad \text{and} \quad |\Psi_-) \in \mathsf{A}_-^\circ \otimes \mathsf{C}_-^\circ. \tag{72}$$

Since the annihilator subspaces are non-overlapping, this contradicts Equation 68. Thus, no such $|\Psi)$ exists for which the set of possible measurement results is described by the PR box pattern $\mathcal{P}$.

5.3 Probabilistic resolutions

We have already noted that we can derive a generalized modal table from a generalized probability table, while respecting the no-signalling principle. Is it possible to do the reverse? That is, if we have a table of possibilities for a modal system, can we find a corresponding table of probabilities? We call this a *probabilistic resolution* of the modal table, and distinguish two different types.

- A *strong probabilistic resolution* assigns zero probability to every impossible result and non-zero probability to every possible result (X).

- A *weak probabilistic resolution* assigns zero probability to every impossible result. However, a "possible" result (X) may be assigned any probability, zero or non-zero. (See the discussion in Subsection 2.1.)

In either case, we require that the resulting table of distributions must satisfy the probabilistic no-signalling principle.

As an example, consider the PR box table $\mathcal{P}$ of Equation 66. It is not difficult to show that this table has only one allowed probabilistic resolution, which is of the strong type:

$$
\begin{array}{cc}
& A^{(2)} \qquad\qquad B^{(2)}
\end{array}
$$

	$A^{(2)}$		$B^{(2)}$	
$A^{(1)}$	$1/2$	0	$1/2$	0
	0	$1/2$	0	$1/2$
$B^{(1)}$	$1/2$	0	0	$1/2$
	0	$1/2$	$1/2$	0

$$\tag{73}$$

Not all general modal tables actually have probabilistic resolutions of either type. Consider the following table (of which we have only shown the relevant parts):

$$\mathcal{N} = (74)$$

The matrix $\mathcal{N}$ has row-blocks $U^{(1)}$, $V^{(1)}$, $W^{(1)}$ and column-blocks $U^{(2)}$, $V^{(2)}$, $W^{(2)}$, with the following 3×3 sub-tables:

$(U^{(1)}, U^{(2)})$:

	X	
		X
X		

$(U^{(1)}, V^{(2)})$:

X		
	X	
		X

$(V^{(1)}, U^{(2)})$:

X		
	X	
		X

$(V^{(1)}, V^{(2)})$:

X		
	X	
		X

$(V^{(1)}, W^{(2)})$:

X		
		X
		X

$(W^{(1)}, V^{(2)})$:

X	X	
		X

$(W^{(1)}, W^{(2)})$:

X		
		X

By inspection, $\mathcal{N}$ satisfies the modal no-signalling principle. When we attempt a probabilistic resolution, we quickly discover that all of the possibilities in the $(U^{(1)}, U^{(2)})$, $(U^{(1)}, V^{(2)})$, $(V^{(1)}, U^{(2)})$ and $(V^{(1)}, V^{(2)})$ sub-tables must be assigned probability $1/3$. We obtain

$$(75)$$

with row-blocks $U^{(1)}$, $V^{(1)}$, $W^{(1)}$ and column-blocks $U^{(2)}$, $V^{(2)}$, $W^{(2)}$:

$(U^{(1)}, U^{(2)})$:

	1/3	
		1/3
1/3		

$(U^{(1)}, V^{(2)})$:

1/3		
	1/3	
		1/3

$(V^{(1)}, U^{(2)})$:

1/3		
	1/3	
		1/3

$(V^{(1)}, V^{(2)})$:

1/3		
	1/3	
		1/3

$(V^{(1)}, W^{(2)})$:

1/3		
		1/3
		1/3

$(W^{(1)}, V^{(2)})$:

1/3	1/3	
		1/3

$(W^{(1)}, W^{(2)})$:

p		
	q	

where we have for clarity omitted zero entries. The trouble arises in the lower-right corner $(W^{(1)}, W^{(2)})$. The probabilistic no-signalling principle imposes two sets of constraints on the probabilities p and q. Comparing to

the $(W^{(1)}, V^{(2)})$ sub-table, we require $p = 2/3$ and $q = 1/3$. Comparing to the $(V^{(1)}, W^{(2)})$ sub-table, we require $p = 1/3$ and $q = 2/3$. We therefore conclude that no probabilistic resolution exists for modal table $\mathcal{N}$.

Under some circumstances, we can guarantee that a probabilistic resolution must exist. Suppose that a general modal table $\mathcal{R}$ arises from a local hidden variable theory. For each particular value h of the hidden variables, the outcomes of all joint measurements are determined. The resulting table $\mathcal{D}_h$ is thus deterministic—that is, each sub-table contains only a single X. The overall table $\mathcal{R}$ is thus a mixture of different $\mathcal{D}_h$ tables. Locality of the hidden variable theory means that each deterministic table $\mathcal{D}_h$ individually satisfies the no-signalling principle.

Suppose there are N distinct deterministic tables $\mathcal{D}_h$. Each deterministic table has an obvious probabilistic resolution in which each X entry is given probability 1. Now we assign each distinct $\mathcal{D}_h$ a probability of $1/N$, and take a mixture of their probabilistic resolutions with these weights. That is, if a particular outcome of a particular joint measurement is possible in M of the deterministic tables, it is assigned an overall probability M/N. The resulting table satisfies the probabilistic no-signalling principle, since it is a convex combination of no-signalling tables. Furthermore, it is a strong probabilistic resolution of $\mathcal{R}$, since it assigns a probability at least $1/N$ to each possible measurement outcome. Therefore, every general modal table arising from a local hidden variable theory has a strong probabilistic resolution.

The converse is certainly false. The PR box table $\mathcal{P}$ of Equation 66 has a strong probabilistic resolution (Equation 73). However, $\mathcal{P}$ is a minimal table, which means it cannot arise as a mixture of deterministic tables that satisfy the no-signalling principle. Therefore $\mathcal{P}$ cannot arise from any local hidden variable theory.

Now consider the modal table $\mathcal{S}$ arising from the $\mathbb{Z}_2$-MQT singlet state, as shown in Equation 65. This has a unique probabilistic resolution, which we display below. Note that some of the possible outcomes have to be assigned probability zero—that is, only a weak probabilistic resolution can be given

for this table:

$$
\begin{array}{c|ccc}
 & X^{(2)} & Y^{(2)} & Z^{(2)} \\
X^{(1)} & \begin{array}{|c|c|} \hline & 1/2 \\ \hline 1/2 & \\ \hline \end{array} & \begin{array}{|c|c|} \hline 1/2 & \\ \hline 0 & 1/2 \\ \hline \end{array} & \begin{array}{|c|c|} \hline 1/2 & 0 \\ \hline & 1/2 \\ \hline \end{array} \\
Y^{(1)} & \begin{array}{|c|c|} \hline 1/2 & 0 \\ \hline & 1/2 \\ \hline \end{array} & \begin{array}{|c|c|} \hline & 1/2 \\ \hline 1/2 & \\ \hline \end{array} & \begin{array}{|c|c|} \hline 1/2 & \\ \hline 0 & 1/2 \\ \hline \end{array} \\
Z^{(1)} & \begin{array}{|c|c|} \hline 1/2 & \\ \hline 0 & 1/2 \\ \hline \end{array} & \begin{array}{|c|c|} \hline 1/2 & 0 \\ \hline & 1/2 \\ \hline \end{array} & \begin{array}{|c|c|} \hline & 1/2 \\ \hline 1/2 & \\ \hline \end{array}
\end{array}
\tag{76}
$$

There are a number of things to remark about the probabilistic resolution in Equation 76. The MQT singlet state $|S)$ gives us an example of a table with a weak probabilistic resolution but not a strong probabilistic resolution. This gives us another proof that the modal properties of $|S)$ (represented in table $\mathcal{S}$) cannot be derived from any local hidden variable theory: if such a theory existed, the table would certainly have a strong probabilistic resolution.

As we have seen, the modal PR box table $\mathcal{P}$ of Equation 66 cannot arise from an entangled composite system in MQT. Nevertheless, the weak probabilistic resolution of Equation 76 does contain a *probabilistic* PR box! Consider the following section of the table:

$$
\begin{array}{c|cc}
 & Z^{(2)} & Y^{(2)} \\
X^{(1)} & \begin{array}{|c|c|} \hline 1/2 & 0 \\ \hline 0 & 1/2 \\ \hline \end{array} & \begin{array}{|c|c|} \hline 1/2 & 0 \\ \hline 0 & 1/2 \\ \hline \end{array} \\
Y^{(1)} & \begin{array}{|c|c|} \hline 1/2 & 0 \\ \hline 0 & 1/2 \\ \hline \end{array} & \begin{array}{|c|c|} \hline 0 & 1/2 \\ \hline 1/2 & 0 \\ \hline \end{array}
\end{array}
\tag{77}
$$

This apparent paradox arises because a weak probabilistic resolution allows probability zero to be assigned to a possible measurement outcome.

A probabilistic PR box cannot arise in actual quantum theory. It follows that the behavior of an entangled composite system in MQT cannot be "simulated" by an entangled composite system in AQT. (This is why the pseudo-telepathy game for $|S)$ described in Subsection 2.3 has no winning strategy if the players can only share entangled states from AQT.)

36

5.4   A hierarchy of modal theories

We have considered several distinct types of two-system modal tables.

- NSP is the set of tables satisfying the no-signalling principle. (This is our "universe" of tables.)

- SPR is the set of tables that have a strong probabilistic resolution.

- WPR is the set of tables that have a weak probabilistic resolution.

- LHV is the set of tables that have a local hidden variable model.

- MQT is the set of tables that can arise from a bipartite system in modal quantum theory.

As we have seen there are several relations between these classes:

$$\text{LHV} \subset \text{SPR} \subset \text{WPR} \subset \text{NSP}. \tag{78}$$

The inclusion relation is strict in each case. The PR box table $\mathcal{P}$ in in Equation 66 is in SPR but not LHV; the $\mathbb{Z}_2$ modal singlet table $\mathcal{S}$ in Equation 65 is in WPR but not SPR; and the table $\mathcal{N}$ in Equation 74 is in NSP but not WPR.

What about the set MQT? It is not hard to see that every table in LHV is also in MQT. We also know there are tables that are in MQT but not in LHV or SPR. Conversely, the PR box $\mathcal{P}$ (Equation 66) is in SPR and WPR but not MQT. It remains to pin down the relation between MQT and WPR. We will prove that MQT $\subset$ WPR—that is, that every table that arises from the state of a bipartite system in MQT must have a weak probabilistic resolution.

To establish this, we will take advantage of several simplifications. Since a weak probabilistic resolution allows us to assign $p = 0$ for some possible outcomes, the addition of possibilities (X entries) to a modal table can never frustrate a weak probabilistic resolution. Therefore, we need only consider minimal modal tables in MQT, those that arise from pure bipartite states.

Every pure bipartite state $|\Psi)$ has a Schmidt decomposition (as in Equation 27) with an integer Schmidt number s. The state vector therefore lies in a subspace we may denote $\mathcal{V} \otimes \mathcal{V}$, with $\dim \mathcal{V} = s$. The space $\mathcal{V}$ is a subspace of the state spaces for the two systems; but we can regard it as the effective state space for the particular situation described by $|\Psi)$. Any measurement

on either subsystem can hence be regarded as a generalized measurement on $\mathcal{V}$. Therefore, we can suppose that $|\Psi)$ is a state of maximum Schmidt number for a pair of identical systems with state spaces $\mathcal{V}$ of dimension s. (The case where $s = 1$ is trivial, so we will assume that $s \geq 2$ and $|\Psi)$ is entangled.)

Generalized measurements whose effect subspaces have dim $\mathsf{E}_a > 1$ can be viewed as "coarse-grained" versions of measurements with one-dimensional ("fine-grained") effects. If we can construct a weak probabilistic resolution for the fine-grained measurements, this will automatically give a resolution for the coarse-grained version. Therefore, we need only consider fine-grained measurements—that is, those whose effect subspaces are one-dimensional.

A fine-grained measurement can be viewed as a spanning set for $\mathcal{V}^*$. Every such spanning set contains a basis, and at least one of these basis effects must be possible for a given state. The "extra" effects can always be assigned probability zero. Therefore, we need only consider basic measurements, those that correspond to basis sets for $\mathcal{V}^*$.

Armed with all of these simplifications, let us consider a pair of identical systems in a pure entangled state $|\Psi^{(12)})$ of maximum Schmidt number. For each pair of basic measurements, we arrive at an $s \times s$ sub-table of possibilities. Let us focus our attention on one such sub-table, with measurement bases $\{(e_j^{(1)}|\}$ (the rows) and $\{(f_k^{(2)}|\}$ (the columns).

For each $(e_j^{(1)}|$, define the set

$$F_j = \{(f_k^{(2)}| : (e_j^{(1)} f_k^{(2)} |\Psi^{(12)}) \neq 0\}. \tag{79}$$

That is, for each system 1 effect, we consider the set of system 2 effects that are jointly possible given state $|\Psi^{(12)})$. Consider next a set E containing d system 1 effects $(e_j^{(1)}|$. Each $(e_j^{(1)}|$ corresponds to a conditional state $|\psi_j^{(2)}) = (e_j^{(1)} |\Psi^{(12)})$. Since $|\Psi^{(12)})$ is maximally entangled, these are non-zero and linearly independent. Hence, the effects in E correspond to a set of system 2 states that span a subspace $\mathsf{M}_E^{(2)}$ of dimension d.

A basic system 2 measurement on M_E must have at least d possible outcomes. These correspond to the system 2 effects in the set $\bigcup_E F_j$. We have shown that the collection $F = \{F_j\}$ of sets has the property that, for any set E of basic system 1 effects,

$$\# \left(\bigcup_E F_j \right) \geq \# (E), \tag{80}$$

38

where $\#(K)$ is the number of elements in finite set K. By Hall's Marriage Theorem [21], we can conclude that the collection F has a set of distinct representatives. That is, for each $\big(e_j^{(1)}\big|$ we can identify a corresponding $\big(f_j^{(2)}\big|$ such that

- $\big(e_j^{(1)}f_j^{(2)}\,\big|\Psi^{(12)}\big) \neq 0$ for all j, and

- $\big(f_i^{(2)}\big| \neq \big(f_j^{(2)}\big|$ when $i \neq j$.

In our sub-table, this means we can identify a set of the possible joint outcomes (the X's) such that each row and each column contains exactly one of them.

We therefore make the following probability assignment. Each impossible joint outcome, of course, is assigned $p = 0$. We also assign $p = 0$ to all of the possible joint outcomes except for those we have identified above, one in each row and column. These are assigned $p = 1/s$.

The same procedure can be applied for each sub-table independently. In every case, the total probability for each row and for each column is $1/s$. Therefore, the probabilistic no-signalling principle is automatically satisfied. Our construction (via Hall's Marriage Theorem) yields a weak probabilistic resolution for the modal table associated with the entangled state $|\Psi^{(12)}\rangle$. Every table that arises from a bipartite state in MQT has a weak probabilistic resolution.

In terms of our hierarchy of modal theories, we have shown that MQT $\subset$ WPR. Our conclusions are summarized in Figure 5.4. It is worth noting that all of the six distinct regions in this diagram are non-empty. Thus, for example, table $\mathcal{S}$ of Equation 65 is in MQT but not SPR; table $\mathcal{P}$ of Equation 66 is in SPR but not MQT; and table $\mathcal{N}$ of Equation 74 is within NSP but not WPR. Other examples are easy to construct.

6 Concluding Remarks

6.1 What MQT has, and what it does not have

As diverting an exercise as MQT is, its real purpose is to shed light on the structure of actual quantum theory. It is remarkable how many of the features of AQT are retained, at least in some form, even in such a primitive theory. An incomplete summary can be found in Figure 6.1. In the left-hand
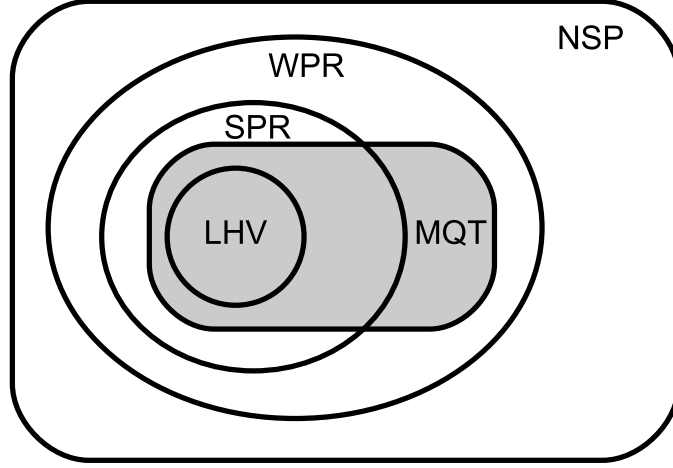
Figure 1: The hierarchy of bipartite states in modal theories.

column we have listed aspects of AQT that are not found in MQT; in the right-hand column, we have listed aspects of AQT that do have analogies in MQT. The key point is that *nothing in the right-hand column logically depends on anything in the left-hand column.*

Furthermore, as we have seen, the process of generalization is very similar in AQT and MQT. In both theories we can develop more general concepts of state, measurement and time evolution, and these generalizations can be characterized in both constructive and axiomatic ways. Both theories can also be extended to more general (probabilistic or modal) theories. Within these more general types of theories, the quantum theories have special properties—e.g., PR boxes are excluded in either theory, and every bipartite state in MQT has a weak probabilistic resolution.

This last point deserves further comment. We have imagined a modal world, one which supports the distinction between "possible" and "impossible" events without necessarily imposing any probability measure. As we have seen, it is not always possible to make a reasonable probability assignment in such a modal world. The table $\mathcal{N}$ of Equation 74 provides an example that respects the modal no-signalling principle, but within which we cannot assign probabilities respecting the probabilistic NSP.

Under what circumstances, then, can we make reasonable probability assignments to a set of possibilities? In the bipartite case, we have shown

MQT does not have:

- Probabilities, expectations

- ($\mathcal{F}$ finite) Continuous sets of states and observables, or continuous time evolution

- Inner product, outer product, orthogonality

- Convexity

- Hermitian conjugation (†)

- Density operators

- Effect operators

- CP maps

- Unextendable product bases

MQT does have:

- "Classical" versus "quantum" theories

- Superposition, interference

- Complementary measurements

- Entanglement

- No local hidden variables

- Kochen-Specker theorem, "free will" theorem

- Superdense coding, teleportation, "steering" of mixtures

- Mixed states, generalized effects, generalized evolution maps

- No cloning theorem

- Nonclassical models of computation

Figure 2: Properties and structures of actual quantum theory that either are or are not present in MQT.

that this can always be done for joint measurements on a modal quantum system. That is, the underlying structure of MQT somehow "makes room" for probabilities. It remains to be seen whether this sheds any light on the way in which probabilities arise in the real world.

6.2 Open problems

Modal quantum theory is an exceptionally rich "toy model" of physics. Despite the known features of the theory summarized in Figure 6.1, there remain many open questions.

- Although we have shown that bipartite systems in MQT support weak probabilistic resolutions, we do not know whether this is true for entangled states of three or more systems.

- We have established many properties of pure entangled states for MQT system, but we know much less about mixed entangled states. For example, we do not know whether there are "bound" entangled states in MQT [22]. (The usual AQT construction cannot be adapted to MQT, since there are no unextendable product bases in MQT.)

- Many results and ideas of quantum information and quantum computation have direct analogues in MQT. For instance, MQT supports both superdense coding and teleportation [6]. It is straightforward to show that the Deutsch-Jozsa oracle algorithm (distinguishing constant and balanced functions with a single query) can be implemented without change on a modal quantum computer with $\mathcal{F} = \mathbb{Z}_3$ [23]. However, a great deal of work remains to be done along these lines.[7]

- It is possible to regard actual quantum theory as a special type of modal quantum theory in which $\mathcal{F} = \mathbb{C}$ and we have special restrictions on the allowed measurements and time evolution operators. What (if anything) can be gained by analyzing AQT in this way?

We believe that the investigation of these and other open problems MQT will shed further light on the mathematical structure of quantum theory.

---

[7]Some observations are obvious. In a world without probabilities, we are interested in the *zero-error capacities* of communication channels and computer algorithms that reach *deterministic* results.

7 Acknowledgments

Appendix

Here we fill in the details of the argument in Subsection 5.2. For convenience, we will suppose that modal quantum systems 1 and 2 are both described by the state space $\mathcal{V}$, and that the same two-outcome measurement is performed on each. The effect subspaces E and F in $\mathcal{V}^*$ are non-overlapping, so that $\mathsf{E} \cap \mathsf{F} = \langle 0 \rangle$. Finally, we assume that the joint possibility table for the state $|\Psi)$ is as follows:

$$
\begin{array}{c c}
 & \begin{array}{cc} \mathsf{E} & \mathsf{F} \end{array} \\
\begin{array}{c} \mathsf{E} \\ \mathsf{F} \end{array} & \begin{array}{|c|c|} \hline \mathsf{X} & \\ \hline & \mathsf{X} \\ \hline \end{array}
\end{array}
\tag{81}
$$

(Each sub-table of Equation 66 is of this form.)

We can find a basis for $\mathcal{V}^*$ of the form $\{(e_i|, (f_m|\}$, where the $\{(e_i|\}$ spans E and $\{(f_m|\}$ spans F. The dual basis $\{|e_i), |f_m)\}$ of $\mathcal{V}$ therefore has the property that $|e_i)$ is annihilated by every $(f_m|$ and $|f_m)$ is annihilated by every $(e_i|$. In fact, $\{|e_i)\}$ spans the annihilator F° and $\{|f_m)\}$ spans E°. We can expand the composite state $|\Psi)$ in this way:

$$
|\Psi) = \sum_{ij} \alpha_{ij} |e_i e_j) + \sum_{in} \beta_{in} |e_i f_n) + \sum_{mj} \gamma_{mj} |f_m e_j) + \sum_{mn} \delta_{mn} |f_m f_n). \tag{82}
$$

From Equation 81, we can see that the effect $\mathsf{E} \otimes \mathsf{F}$ is impossible, which implies that $(e_i f_n | \Psi) = \beta_{in} = 0$ for every i, n. In the same way, because $\mathsf{F} \otimes \mathsf{E}$ is impossible, $\gamma_{mj} = 0$ for every m, j. Therefore,

$$|\Psi) = |\Psi_{ee}) + |\Psi_{ff}), \qquad (83)$$

where $|\Psi_{ee}) \in \mathsf{F}^{\circ} \otimes \mathsf{F}^{\circ}$ and $|\Psi_{ff}) \in \mathsf{E}^{\circ} \otimes \mathsf{E}^{\circ}$.

Though we have supposed that the two systems are of the same type and that the same measurement is made on each, it is easy to adapt this argument to more general situations, provided the effect subspaces are non-overlapping.

References

[1] Robert Spekkens, "Evidence for the epistemic view of quantum states: A toy theory", *Phys. Rev. A* **75**, 032110 (2007).

[2] J. S. Bell, "On the Einstein-Podolsky-Rosen paradox", *Physics* **1**, 195 (1964).

[3] Lucien Hardy, "Reformulating and reconstructing quantum theory", e-print arxiv: 1104.2066.

[4] G. Chiribella, G.M. DAriano and P. Perinotti, "Informational derivation of quantum theory", *Physical Review A* **84**, 012311-1 (2011).

[5] Lus Masanes and Markus P. Mller, "A derivation of quantum theory from physical requirements", *New J. Phys.* **13**, 063001 (2011).

[6] B. Schumacher and M. D. Westmoreland, "Modal quantum theory", e-print arXiv:1010:2937.

[7] G. Hughes and M. Cresswell, *A New Introduction to Modal Logic* (Routledge, London, 1996).

[8] B. Schumacher and M. D. Westmoreland, "Non-contextuality and free will in modal quantum theory", e-print arXiv:1010:5452.

[9] Christoph Simon, Vladimìr Bužek and Nicolas Gisin, "The no-signaling condition and quantum dynamics", *Phys. Rev. Lett.* **87**, 170405 (2001).

[10] S. Kochen and E. Specker, "The problem of hidden variables in quantum mechanics", *J. Math. Mech.* **17**, 59–88 (1967).

[11] B. Schumacher and M. Westmoreland, "Locality Non-contextuality and Free Will in Modal Quantum Theory", poster presented at Conceptual Foundations and Foils for QIP, Perimeter Institute (2011). URL: http://www.perimeterinstitute.ca/Events/

[12] Gilles Brassard, Anne Broadbent, Alain Tapp, "Multi-Party Pseudo-Telepathy", *Proceedings of the 8th International Workshop on Algorithms and Data Structures, Lecture Notes in Computer Science* **2748**, 1-11, (2003).

[13] B. Schumacher and M. D. Westmoreland, *Quantum Processes, Systems and Information* (Cambridge University Press, Cambridge, 2010).

[14] Paul Halmos, *Naive Set Theory* (Springer - Verlag, New York, 1974).

[15] I. M. Gelfand and M. A. Neumark, "On the embedding of normed rings into the ring of operators in Hilbert space", *Rec. Math. [Mat. Sbornik] N.S.* **54**, 197213 (1943).

[16] Hughston, L. P., Jozsa, R., and Wootters, W. K., "A complete classification of quantum ensembles having a given density matrix", *Phys. Lett. A* **183**, 1418 (1993).

[17] Schrödinger, E. (1936). "Probability relations between separated systems", *Proc. Camb. Phil. Soc. 32*, 446452.

[18] Jaynes, E. T., "Information theory and statistical mechanics. II",*Phys. Rev.* **108**(2), 171190 (1957).

[19] H. M. Wiseman, S. J. Jones, and A. C. Doherty, "Steering, Entanglement, Nonlocality, and the Einstein-Podolsky-Rosen Paradox", *Phys. Rev. Lett.* **98**, 140402 (2007).

[20] S. Popescu and D. Rohrlich, "Nonlocality as an axiom", *Found. Phys.* **24**, 379–385 (1994).

[21] P. Hall, "On representatives of subsets", *J. London Math. Soc.* **10**, 26-30 (1935).

[22] C.H. Bennett, D.P. DiVincenzo, T. Mor, P.W. Shor, J.A. Smolin, B.M. Terhal, "Unextendible Product Bases and Bound Entanglement", *Phys.Rev.Lett.* **82**, 5385 (1999).

[23] David Deutsch and Richard Jozsa, "Rapid solutions of problems by quantum computation", *Proceedings of the Royal Society of London A* **439**, 553 (1992).